

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE LETRAS
PROGRAMA DE PÓS GRADUAÇÃO EM LINGUÍSTICA

Bianca Rossini Fernandes

A aprendizagem implícita de dependências não adjacentes por crianças de 7 anos

Juiz de Fora

2025

Bianca Rossini Fernandes

A aprendizagem implícita de dependências não adjacentes por crianças de 7 anos

Dissertação apresentada ao Programa de Pós-Graduação em Linguística da Faculdade de Letras, da Universidade Federal de Juiz de Fora como requisito parcial para a obtenção do título de Mestre em Linguística. Área de concentração: Linguística e Cognição.

Orientadora: Profa. Dr^a Cristina Name

Juiz de Fora

2025

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Fernandes, Bianca Rossini .

A aprendizagem implícita de dependências não adjacentes por crianças de 7 anos / Bianca Rossini Fernandes. -- 2025.
91 f. : il.

Orientadora: Cristina Name

Dissertação (mestrado acadêmico) - Universidade Federal de Juiz de Fora, Faculdade de Letras. Programa de Pós-Graduação em Linguística, 2025.

1. dependências não adjacentes. 2. aprendizagem estatística. 3. crianças de 7 anos de idade. I. Name, Cristina , orient. II. Título.

Bianca Rossini Fernandes

A aprendizagem implícita de dependências não-adjacentes por crianças de 7 anos

Dissertação
apresentada ao
Programa de Pós-
Graduação em
Linguística
da Universidade
Federal de Juiz de
Fora, como requisito
parcial à obtenção do
título de mestre em
linguística. Área de
concentração:
Linguística.

Aprovada em 10 de março de 2025.

BANCA EXAMINADORA

Profa. Dra. Maria Cristina Lobo Name - Orientadora
Universidade Federal de Juiz de Fora

Profa. Dra. Mercedes Marcilese
Universidade Federal de Juiz de Fora

Prof. Dr. Cândido Samuel Fonseca de Oliveira
Centro Federal de Educação Tecnológica de Minas Gerais

Profa. Dra. Ana Paula Martins Alves Salgado
Universidade Federal Rural da Amazônia

Juiz de Fora, 06/03/2025.



Documento assinado eletronicamente por **Maria Cristina Lobo Name, Professor(a)**, em 10/03/2025, às 18:42, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Mercedes Marcilese, Professor(a)**, em 11/03/2025, às 09:15, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ana Paula Martins Alves Salgado, Usuário Externo**, em 11/03/2025, às 10:35, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Cândido Samuel Fonseca de Oliveira, Usuário Externo**, em 11/03/2025, às 17:11, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **2278112** e o código CRC **3D15BE2A**.

AGRADECIMENTOS

Sempre ouvi dizer que o caminho para a pós-graduação não é fácil. E realmente não é. Mas me considero uma pessoa de sorte por ter podido contar com pessoas que tornaram essa jornada muito mais fascinante do que penosa.

Começo agradecendo à minha orientadora, Cristina, que sabe conciliar profissionalismo com humanidade e paciência, de maneira que, ao mesmo tempo em que sempre me deu muita autonomia, esteve por perto para me auxiliar e ensinar. Sei que, toda vez que lhe agradeço pessoalmente, você me diz que está apenas fazendo seu trabalho, mas gostaria que soubesse que faz muito mais do que isso. Então, muito obrigada, mais uma vez, por esses dois anos de tanto aprendizado.

Agradeço também à minha inspiração de vida, Lily, minha filha amada, que sempre entendeu que a mamãe às vezes precisava estudar nos fins de semana, feriados, férias...

Ao meu marido, Deco, companheiro de vida, que ouviu incansavelmente sobre assuntos de que não entendia, que assistiu às minhas apresentações, que me auxiliou de todas as maneiras possíveis e em quem sempre pude contar com apoio incondicional.

À minha mãe, ao meu padrasto e à minha avó, que, mesmo sem ter ideia do que eu estava estudando, sempre tiveram fé de que eu conseguiria.

Aos professores Cândido Oliveira e Ana Paula Salgado, por terem sido extremamente generosos durante meu exame de qualificação, contribuindo com direcionamentos muito importantes para este estudo. Obrigada, novamente, por aceitarem fazer parte da banca da minha defesa.

À professora Mercedes Marcilese, que foi minha primeira professora na pós, quando ainda eu cursava disciplina isolada, e que me apresentou à Psicolinguística. Estou muito grata por tê-la como membro da minha banca.

Aos professores do PPG, que contribuíram imensamente para minha formação. À secretária do PPG, em especial à Isabella, que sempre se mostrou disponível para me ajudar com as inevitáveis burocracias.

Às crianças e seus responsáveis, que aceitaram fazer parte deste projeto. A troca com todos foi muito especial, além de ter proporcionado momentos que ficarão para sempre na minha memória.

À UFJF, por proporcionar um curso de tamanha qualidade, e à CAPES, pelo apoio financeiro, que foi essencial para que eu pudesse me dedicar a esta pesquisa com tranquilidade.

RESUMO

O presente estudo investiga a aprendizagem implícita de dependências não adjacentes (DNAs) por crianças brasileiras de 7 anos de idade. As DNAs são caracterizadas por estabelecer uma relação de codependência entre dois elementos da estrutura linguística com um material linguístico interveniente variável (ex., minha caneta, essa mesa; ele está dormindo, ela está comendo). Assim, estabelecem-se diversas DNAs nas línguas naturais combinando diferentes classes ou categorias funcionais, que, por sua vez, possuem diferentes itens que estabelecem essa relação de codependência. Desse modo, o processamento e aprendizagem das DNAs são imprescindíveis no processo de aquisição de linguagem. Especificamente, focalizamos a faixa etária dos 7 anos de idade, que passa por transformações cognitivas e metacognitivas e por ainda levantar certas questões em relação à aprendizagem implícita de DNAs. São objetivos (i) verificar, a partir de uma tarefa experimental, se crianças aos sete anos de idade são sensíveis às DNAs contidas em conjuntos de uma classe morfológica de uma língua pseudonatural, mesmo que por um tempo curto de exposição; (ii) investigar em que medida a atenção direcionada aos padrões linguísticos apresentados (DNAs) é requisito para sua aprendizagem implícita nessa faixa etária; e (iii) avaliar a eficácia a Tarefa de Escolha Forçada como um método experimental na investigação de aprendizagem de DNAs por crianças de 7 anos de idade, analisando possíveis necessidades de modificações e adaptações. Foi conduzido um estudo experimental com 32 crianças. Desse total, 16 participaram do grupo experimental, onde foram expostas a estímulos linguísticos contendo DNAs probabilísticas durante a fase de familiarização, enquanto a outra metade formou o grupo controle, no qual os participantes ouviram sons instrumentais durante a mesma fase. Os resultados não revelaram diferença estatisticamente significativa nas taxas de acertos entre os grupos experimental e controle. Além disso, a taxa de acerto dos participantes do grupo experimental se mostrou um pouco acima do nível da chance, sugerindo que as crianças do grupo experimental não foram sensíveis aos padrões de DNAs apresentados. Quando comparados aos dados desta pesquisa com dados de bebês e adultos, essa falta de sensibilidade pode ser atribuída a mudanças nos processos e mecanismos de aprendizagem, combinadas com a necessidade de uma tarefa experimental comportamental adequada para essa faixa etária.

Palavras-chave: dependências não adjacentes; crianças de 7 anos de idade; aprendizagem implícita; tarefa de Escolha Forçada

ABSTRACT

The current study investigates the implicit learning of non-adjacent dependencies (NADs) by 7-year-old Brazilian children. NADs establish a codependent relationship between two elements with a variable intervening material within a language structure (e.g., minha caneta, essa mesa; ele está dormindo, ela está comendo). These structures are commonly found in natural language, combining different functional classes or categories. These classes or categories typically include items that form codependent relationships, making the processing and learning of NADs essential for language acquisition. Specifically, this study focuses on 7-year-old children, a developmental stage marked by cognitive and metacognitive transformations. Besides, questions regarding the implicit learning of NADs at this age remain unresolved. The objectives were (i) to verify, through a behavioural experimental task, whether 7-year-old children were sensitive to NADs embedded in a morphological class set of a pseudonatural language; (ii) to investigate the extent to which directed attention to the language patterns plays a role in its implicit learning; (iii) to evaluate the efficacy of the Two-Forced Alternative Choice task (2AFC) as an experimental method for studying NAD learning in 7 year-olds, identifying potential needs for modification and adaptation. Thirty-two children participated in the study. Of this total, 16 were part of the experimental group and were exposed to linguistic stimuli containing probabilistic NADs during the familiarization phase. The remaining 16 children formed the control group, which had no contact with linguistic stimuli during the same phase. The results did not reveal a statistically significant difference in accuracy rates between the experimental and control groups. Besides, the accuracy rate among experimental group participants was marginally above chance level. This suggests that the children in the experimental group were not sensitive to the presented NAD patterns. When compared to data from infants and adults, this lack of sensitivity may be attributed to developmental changes in learning processes and mechanisms, combined with the need of an appropriately designed behavioral experimental task for this age group.

Keywords: Non-adjacent dependencies; 7-year-olds; implicit learning; two-forced alternative task.

LISTA DE ILUSTRAÇÕES

Figura 1	- Tarefa experimental Monitoramento de Palavras (Lammertink, 2019) ..	51
Figura 2	- <i>Slides</i> do teste de leitura	59
Figura 3	- Tela apresentada aos participantes antes do início da atividade	60
Figura 4	- Exemplo das imagens durante a fase de familiarização	61
Figura 5	- Exemplo do <i>layout</i> da tela durante a fase de teste	62
Quadro 1	- Exemplos de padrões dos estímulos acústicos experimentais	66
Gráfico 1	- Acertos e erros cometidos pelos participantes do grupo experimental durante a fase teste	68
Apêndice A	- Questionário sobre as crianças em relação às suas habilidades e hábitos de leitura	85
Apêndice B	- Folha A4 em que as crianças escolhiam o nome para o robô do experimento	88
Anexo A	- Palavras novas usadas na fase de teste para todos os participantes	89
Anexo B	- Estímulos sonoros da fase de familiarização para os participantes do grupo experimental	90

LISTA DE TABELAS

Tabela 1 - Média de acertos e erros pelos participantes de ambos os grupos durante a fase teste	68
Tabela 2 - Tabela de contingência relacionando a acurácia das respostas pelos participantes do grupo experimental em relação às suas respectivas idades	70
Tabela 3 - Taxa de acurácia das respostas pelos participantes do grupo experimental em relação ao tempo de familiarização (máximo de 3 minutos)	71
Tabela 4 - Média de desempenho dos participantes do grupo experimental em relação ao sexo	72
Tabela 5 - Média de acertos dos participantes do grupo experimental em relação ao ano escolar	72
Tabela 6 - Médias, tempos mínimos e máximos durante a fase de familiarização para ambos os grupos	74
Tabela 7 - Teste χ^2 para o tempo de familiarização em ambos os grupos	74
Tabela 8 - Comparação entre os resultados obtidos no presente trabalho e os resultados do estudo de base	75

SUMÁRIO

1 INTRODUÇÃO.....	11
2 DNAs: O QUE SÃO E COMO AS APRENDEMOS.....	16
2.1 DNAs: CARACTERÍSTICAS E EXEMPLOS.....	16
2.2 APRENDIZAGEM ESTATÍSTICA E HABILIDADE COMPUTACIONAL NA APRENDIZAGEM DE DNAs.....	19
2.3 A APRENDIZAGEM DE DNAs.....	24
2.3.1 Pistas fonológicas: pausas e fronteiras como facilitadores.....	25
2.3.2 Variabilidade do X.....	27
2.3.3 A aprendizagem de DAs e DNAs.....	28
2.3.4 Condições favoráveis para a aprendizagem de DNAs.....	28
2.4 CONCLUSÃO.....	30
3 O PERFIL COGNITIVO E METACOGNITIVO DE CRIANÇAS DE 7 ANOS.....	32
3.1 PROCESSOS COGNITIVOS E METACOGNITIVOS NA INFÂNCIA: O PAPEL DA IDADE NA APRENDIZAGEM.....	32
3.2 CONCLUSÃO.....	40
4 DNAs DETERMINÍSTICAS E PROBABILÍSTICAS, O PAPEL DA ATENÇÃO NA APRENDIZAGEM DE DNAs E A INVESTIGAÇÃO DE APRENDIZAGEM IMPLÍCITA.....	41
4.1 PADRÕES DETERMINÍSTICOS E PROBABILÍSTICOS DE DNAs.....	42
4.2 O PAPEL DA ATENÇÃO NA INVESTIGAÇÃO DE APRENDIZAGEM DE DNAs..	45
4.3 A INVESTIGAÇÃO DE APRENDIZAGEM IMPLÍCITA EM TAREFAS EXPERIMENTAIS.....	49
4.3.1 Discussão dos Resultados e Análises dos Estudos de Base.....	52
3.5 CONCLUSÃO.....	54
5 ATIVIDADE EXPERIMENTAL.....	56
5.1 A TAREFA DE ESCOLHA FORÇADA.....	56
5.2 EXPERIMENTO.....	58
5.2.1 Objetivo.....	62
5.2.2 Método.....	62
5.2.2.1 Participantes.....	64
5.2.2.2 Materiais utilizados.....	65
5.2.2.3 O ambiente.....	66
5.2.3 Análise e discussão.....	67
5.2.3.1 Resultados.....	67
5.2.3.2 Análises Exploratórias.....	69
5.2.3.3 Tempo de familiarização para os grupos experimental e controle.....	73
5.2.3.4 Discussão.....	74
5.3 CONCLUSÃO.....	77

6. CONSIDERAÇÕES FINAIS.....	79
REFERÊNCIAS.....	81
APÊNDICES.....	85
ANEXOS.....	89

1 INTRODUÇÃO

Esta pesquisa investiga a aprendizagem implícita de dependências não adjacentes (doravante, DNAs) em uma língua pseudonatural por crianças de 7 anos de idade. Por ser uma faixa etária ainda pouco estudada, sobre a qual há ainda poucos dados disponíveis em relação à aprendizagem de DNAs, além de ser uma idade que passa por transformações cognitivas e metacognitivas (ex. maturação do córtex pré-frontal, desenvolvimento de funções executivas como controle inibitório e memória de trabalho, etc), esse trabalho propõe responder algumas questões deixadas em aberto por estudos que foram os pioneiros em investigar o comportamento de participantes crianças nesse tipo de aprendizagem (Lammertink et al., 2019; Marimon et al., 2021).

Entende-se que DNAs em uma língua natural são relações de codependências entre dois materiais linguísticos que se distam e são separados por um outro material interveniente variável, expressando relações de concordância (de gênero, número, pessoa, etc), como por exemplo: *aquela **estrela**, métodos **variáveis**, ela **está comendo***. Os elementos que compõem as DNAs em línguas naturais, normalmente, ocorrem em classes ou conjuntos de itens funcionais. No entanto, esses padrões podem se diferenciar de uma língua para outra, o que veremos com mais detalhes na seção 2.1 deste trabalho.

As DNAs, na maioria das vezes, são relações estabelecidas morfossintaticamente, sendo parte intrínseca das línguas naturais, de modo que a identificação, a abstração e generalização dos padrões de DNAs são fundamentais no processo de aquisição de linguagem, exigindo processamento linguísticos e estatísticos complexos (Laguardia, 2016). A configuração utilizada para caracterizar uma DNA em estudos sobre o tema é aXb , onde a é o primeiro elemento constituinte da relação de codependência, X é o material interveniente e b é o elemento final dessa codependência.

Aqui, assumimos que a aprendizagem estatística em relação à aquisição de padrões de DNAs tem o papel de auxiliar o indivíduo a perceber e extrair informações de regularidades que esses padrões apresentam (Saffran et al, 1996; Aslin, Saffran e Newport, 1999; Gómez, 2002; Newport e Aslin, 2004), o que a literatura aponta tanto por bebês quanto por adultos (Gómez, 2002; Gómez e Maye, 2005; Name et al., 2011; Laguardia et al., 2013; Name e Shi, 2015; Laguardia et al., 2015; Henrique, 2023). Teremos uma breve revisão sobre a aprendizagem de DNAs em bebês e adultos na seção 2.3, a fim de sistematizar algumas descobertas sobre o tema.

Além da aprendizagem estatística, também assumimos que outros mecanismos de aprendizagem estão envolvidos nesse tipo de aquisição. Mecanismos esses que podem pertencer a domínios mais gerais, como a atenção, por exemplo, assim como mecanismos mais especializados, incluindo uma habilidade computacional linguística. Essa habilidade computacional seria uma das ferramentas disponíveis para a abstração e generalização de regras subjacentes aos elementos compilados nesse tratamento estatístico, regras nas quais os padrões de DNAs estariam inseridos (Marcus et al. 1999; Peña et al., 2002). Sobre esse conceito e debate na literatura, teremos mais detalhes nas seções 2.2 e 2.3.

Em estudos sobre aprendizagem de DNAs, são utilizados os seguintes padrões de DNAs: o determinístico e o probabilístico, sendo o primeiro uma relação de dependência de 1 para 1 (se tenho *a*, é certo que terei *b*), e o último assumindo a possibilidade de um elemento se concatenar a diferentes elementos de um dado conjunto ou categoria (se tenho um elemento *a* que pertence ao conjunto A, terei algum elemento *b*, pertencente ao conjunto B), como por exemplo, um elemento pertencente à classe dos determinantes combinando-se com outro elemento que pertença à uma classe de vogal temática (*um* camelo, *aquele* pente, *meu* sorriso). Veremos com mais detalhes, na seção 4.1, a denominação e uso desses dois padrões.

Algumas pesquisas buscaram investigar em que medida se dá a aprendizagem implícita de DNAs em diferentes faixas etárias. Os resultados de pesquisas que investigaram a aprendizagem implícita de DNAs sugerem que bebês já aos 11 meses de idade (Name et al., 2011; Laguardia et al., 2013; Name e Shi, 2015; Laguardia et al., 2015) e adultos (Gómez, 2002; Henrique, 2023) são sensíveis a esses padrões de codependências linguísticas.

Uma das definições de aprendizagem implícita caracteriza a situação quando o indivíduo não está ciente de que se está aprendendo algo ao executar ou participar de uma tarefa que envolva alguma aprendizagem. Assim, mesmo que o indivíduo não tenha consciência de que aprendeu algo, essa aprendizagem tem influência relevante em situações subsequentes, mesmo que os indivíduos tenham sido expostos aos estímulos por um curto período de tempo e que a resposta a esses estímulos pareça ser de caráter intuitivo (Cleeremans, 2009).

Neste estudo, investigamos a aprendizagem implícita de DNAs de padrões probabilísticos em uma língua pseudonatural por crianças de 7 anos. Para a idade investigada nesta pesquisa - crianças de 7 anos de idade - existem particularidades a serem consideradas. Como veremos na seção 3.1, há estudos que sugerem que a criança nessa idade passa por um processo de mudança de estratégias na aprendizagem (Rabi e Minda, 2014; Skeide e Friederici, 2016; Roark e Holt, 2019; Schaadt et al., 2019; Van Der Kant et al., 2020, Paul et

al., 2021), em decorrência da maturação do córtex pré-frontal (Skeide e Friederici, 2016). Em contrapartida, há trabalhos (para uma revisão, ver Thiessen et al., 2016) que defendem que, na verdade, alguns mecanismos perdem eficiência ao longo dos anos, devido ao desenvolvimento em outros aspectos, como experiência linguística e funções do controle executivo, por exemplo. Além disso, há evidência de que processos metacognitivos também passem por mudanças durante essa fase (Bialystok, 1986; Edwards & Kirkpatrick, 1999). Esses processos metacognitivos parecem caminhar juntos com a trajetória pedagógica da criança de 7 anos na educação formal (Barrera e Maluf, 2003), quando ela passa pelo processo de alfabetização, aprimorando escrita e leitura.

Assim, a partir de dois trabalhos que investigaram a aprendizagem de DNAs por crianças de 7 anos (Lammertink et al., 2019; Marimon et al., 2021), foram levantadas algumas questões que serão exploradas neste trabalho.

A primeira delas é que esses estudos na faixa etária dos 7 anos só se voltaram para a aprendizagem implícita de padrões de DNAs determinísticos, mas não probabilísticos. A segunda questão é que ainda não se pode afirmar que a atenção dos participantes aos estímulos seja crucial para a investigação desse tipo de aquisição em crianças ainda em estágios de desenvolvimento cognitivo e metalinguístico. E por último, mas não menos importante, é a questão da aplicação da tarefa experimental de Escolha Forçada, considerada *offline* e que consiste em apresentar duas opções para o participante, para que este escolha a melhor opção de acordo com a instrução dada. Esse tipo de tarefa pode ser aplicado tanto para investigação de aquisição de linguagem/aprendizagem de línguas ou em outros contextos não linguísticos. Os autores desses trabalhos (Lammertink et al., 2019; Marimon et al., 2021), diante dos resultados encontrados a partir desse tipo de tarefa, sugerem que a tarefa de Escolha Forçada não seja sensível o suficiente para apontar informações mais claras sobre a aprendizagem implícita de padrões de DNAs nessa faixa etária investigada neste trabalho.

Diante de tudo isso, ainda não é claro se crianças de 7 anos, que já possuem uma língua materna consolidada, aprendem, de maneira implícita, padrões de DNAs de uma outra língua, regidos por regras combinatórias de elementos inseridos em categorias funcionais. Além disso, questionamos se a atenção ao padrão linguístico é relevante para essa aprendizagem e se a tarefa de Escolha Forçada é, de fato, adequada para investir a aprendizagem implícita de DNAs nessa faixa etária.

Com isso, buscamos investigar essas lacunas deixadas por estudos anteriores a fim de contribuir para o estudo acerca da aprendizagem de DNAs por crianças aos 7 anos de idade, seja na sua própria língua ou em uma segunda língua.

Visto que há evidências na literatura de aprendizagem de padrões de DNAs probabilísticos, tanto em bebês (Name et al., 2011; Laguardia et al., 2013; Name e Shi, 2015; Laguardia et al., 2015) quanto em adultos (Van der Bos et al., 2012; Henrique, 2023), nossa hipótese de trabalho aqui adotada é que crianças de 7 anos de idade também são capazes de aprender implicitamente padrões de DNAs probabilísticos.

O objetivo geral deste trabalho é investigar a sensibilidade de crianças de sete anos a padrões de DNAs probabilísticos contidos em conjuntos de classes morfológicas de uma língua pseudonatural, considerando a relação entre a atenção aos estímulos e a aprendizagem implícita.

Assim, os objetivos específicos são:

(1) observar, a partir de uma tarefa experimental, se crianças aos sete anos de idade são sensíveis às DNAs contidas em conjuntos de uma classe morfológica de uma língua pseudonatural, mesmo que por um tempo curto de exposição;

(2) investigar em que medida a atenção direcionada aos padrões linguísticos apresentados (DNAs) é requisito para a aprendizagem implícita nessa faixa etária;

(3) e avaliar a eficácia da Tarefa de Escolha Forçada como uma atividade experimental adequada para a investigação de aprendizagem de DNAs por crianças de 7 anos de idade, analisando possíveis necessidades de modificações e adaptações.

Esta dissertação conta com a seguinte estrutura:

Os três primeiros capítulos são de fundamentação teórica.

No capítulo 2, apresentaremos as dependências não adjacentes e estudos sobre sua aprendizagem, além da noção e fundamentos sobre a aprendizagem estatística e habilidade linguística computacional.

Em seguida, no capítulo 3, nos aprofundaremos em estudos sobre aprendizagem na faixa etária aqui investigada. Apresentaremos estudos sobre a aprendizagem de DNAs por crianças de 7 anos, além de pesquisas que olharam para outros aspectos sobre essa mesma idade, como por exemplo, segmentação de palavras, aprendizagem implícita, categorização, consciência metalinguística, etc.

No capítulo 4, através também de uma fundamentação teórica, traremos trabalhos que dialogam diretamente com as questões suscitadas neste trabalho, como os padrões de DNAs determinísticos e probabilísticos, o papel da atenção na aprendizagem de DNAs e a investigação de aprendizagem implícita em trabalhos empíricos.

No capítulo 5, descreveremos detalhadamente a tarefa experimental e todas suas fases, assim como características dos participantes, variáveis independentes e dependente, materiais utilizados, resultados, além de análise e discussão.

Por fim, teremos, no capítulo 6, as considerações finais e encaminhamentos para trabalhos futuros.

2 DNAS: O QUE SÃO E COMO AS APRENDEMOS

Este capítulo propõe-se a revisar aspectos relevantes e a levantar informações sobre as dependências não adjacentes e de alguns dos estudos que foram conduzidos a partir desse tema. Na seção 2.1., teremos uma apresentação do que são as DNAs e algumas possibilidades de ocorrências nas línguas.

Na seção 2.2., apresentaremos as noções de aprendizagem estatística e habilidade linguística computacional. A aprendizagem estatística na literatura tem mostrado relevância em estudos onde os participantes conseguem perceber e extrair informações dos estímulos, de maneira que, a partir de um teste, esses participantes demonstram desempenhos satisfatórios em investigações de aprendizagem de variados padrões (segmentação de palavras, categorização de palavras, dependências adjacentes, dependências não adjacentes, etc.). A habilidade linguística computacional, por outro lado, defende que, possuímos mecanismos especializados para abstração e generalização das regras, que foram percebidas e extraídas em decorrência da aprendizagem estatística. Desse modo, esses dois construtos servirão como apoio teórico para as argumentações da investigação da aprendizagem implícita de DNAs em crianças de 7 anos.

Em seguida, na seção 2.3., teremos uma visão geral de literatura sobre a aprendizagem de DNAs, a fim de sistematizar o que já temos de direcionamento sobre o tema.

2.1 DNAS: CARACTERÍSTICAS E EXEMPLOS

As DNAs são padrões linguísticos que apresentam certas regularidades e são bastante frequentes nas línguas naturais.

Como vimos na Introdução, as DNAs nas línguas naturais, além de estabelecerem relações de codependência, também estabelecem relações de concordância. Essas relações de concordância podem se apresentar na esfera morfológica, de modo que a sentença (1) engloba ambos os tipos e exemplifica como as DNAs estão presentes e são partes intrínsecas das línguas naturais.

- (1) Aquelas casas são vendidas somente para empresários.

Assim, em (1), temos cinco relações de dependências com material interveniente, ou seja, estamos diante de cinco padrões de DNAs no português brasileiro. Os padrões de DNAs são caracterizados em estudos como aXb , onde a é o primeiro elemento constituinte da dependência, X é o material interveniente, e b é o segundo elemento parte da relação de dependência. Dessa forma, podemos sistematizar (1) das seguintes maneiras:

Em (a) e (b), as primeiras relações de codependência são estabelecidas entre o determinante (Det) “*aquela*” - de gênero feminino, seguido pela marca de plural -s - e o nome (N) “*casa*” composto pelo radical *cas-* mais vogal temática -a e a marca -s de plural do nome.

(a) Aquelas casas

(b) Aquelas casas

Na sequência, em (c), a relação está entre a marca de plural do DP “*aquelas casas*” e a desinência verbal de *ser*, também no plural, no Presente do Indicativo.

(c) Aquelas casas são¹

E por fim, em (d) e (e), temos a relação entre o verbo *ser* e a terminação -idas do particípio do verbo *vender*, que por sua vez, também concorda com o DP *aquelas casas* em gênero e número.

(d) são vendidas

(e) Aquelas casas são vendidas

Uma outra característica presente nas DNAs em línguas naturais são as relações de codependência entre conjuntos de itens funcionais. Para ilustrar melhor, tomemos como exemplo a relação entre determinante e vogal temática no Português Brasileiro. Na configuração aXb, *a* pode ser qualquer item da classe dos determinantes (o, seu, esse, etc.), *X* será o radical do nome em questão (bonec-, espelh-, present-, etc.) e, *b* será a vogal temática concatenada ao radical do nome *X* (o, e). O que podemos perceber a partir desse exemplo é que, apesar de haver a codependência entre os elementos *a* e *b* (ex., o boneco), existe uma possibilidade de variação entre os elementos envolvidos nesse padrão linguístico, ou seja, *a* podendo ser *A*₁, *A*₂, *A*₃, etc (ex., o boneco, seu boneco, esse boneco) e *b* podendo ser *B*₁, *B*₂, *B*₃, etc (ex., o/seu/esse boneco, presentee, planetaa), com um material *X* interveniente de variação incalculável. Em um artigo de revisão, Wilson et al. (2020) classificaram essas relações de DNAs como classes perceptuais, onde é possível fazer generalizações entre todos os elementos contidos em cada classe a partir de uma regra aprendida; esses padrões de DNAs também são conhecidos como probabilísticos. Na seção 4.1, aprofundaremos melhor em como esses padrões probabilísticos são investigados na literatura.

Outro aspecto importante a ser ressaltado é que, embora as DNAs sejam bastante comuns nas línguas naturais, os elementos que as constituem são distintos entre as línguas. Se

¹ Aqui, apesar de não parecer que há material interveniente, a relação de dependência se inicia bem antes do elemento final de concordância (*são*). Por isso, embora *casas* e *são* sejam termos adjacentes fonologicamente, de acordo com uma análise morfossintática a partir do Programa Minimalista de Chomsky (2000), juntos eles configuram uma relação não adjacente, ou seja, “*casas*” e “*são*” não sofrem operação de *merge*.

compararmos o inglês com o português, por exemplo, não teremos relações de dependência de gênero entre determinantes e nomes, somente de número. Além disso, as relações de codependência entre Det e N podem se apresentar como uma alteração fonética em N (ex: *a* por *e*) ou uma adição da marca *-s* ao final (*The man, the woman, these men, these women, the book, these books*). No entanto, a ideia de que essas relações de dependências estão contidas em conjuntos também se aplica (*the, these, those* são itens de um mesmo conjunto).

Uma outra possibilidade de relação de dependência não adjacente vem através de exemplos no inglês em um sintagma verbal (em inglês, *Verbal Phrase* ou *VP*), onde um mesmo conjunto de itens funcionais pode configurar mais de um tipo de relações de codependências. Observe os exemplos:

Em (2), a relação de codependência é estabelecida entre o verbo auxiliar *to be* (*is, are, was, were, etc.*) e a forma do particípio passado para verbos regulares *-ed*.

(2) *The women are threatened.*

Em (3), a relação é entre o verbo auxiliar *have* (podendo também assumir a forma *has* ou *had*) e a forma do particípio passado para verbos regulares *-ed*.

(3) *The women have threatened the system.*

E em (4), a DNA é estabelecida através do verbo auxiliar *to be* (*is, are, was, were, etc.*) e a terminação verbal *-ing*.

(4) *The women are threatening the system.*

Embora cada um desses exemplos sejam aplicados em contextos semânticos distintos, eles ilustram uma outra possível característica das DNAs presentes em línguas naturais: além da relação de codependência entre conjuntos, um mesmo conjunto (ex., verbo auxiliar *to be*), pode estabelecer dependências distintas com elementos diferentes que não pertencem ao mesmo conjunto (ex., *-ed* e *-ing*); ou ainda um elemento *b* (ex. terminação *-ed*) pode estabelecer relações de codependências com conjuntos distintos de *a* (ex. verbo auxiliar *to be* ou verbo auxiliar *have*). Assim, é importante compreender que há especificidades em cada língua e diferentes tipos de relações entre itens não adjacentes que envolvem aspectos morfossintáticos de diversas classes e estruturas.

Esta seção não tem o intuito de descrever exhaustivamente todos os tipos de relações que os padrões de DNAs estabelecem. No entanto, mostramos como essas relações que as DNAs apresentam nas línguas naturais estão longe de serem simples regularidades. Dessa forma, diversos estudos vêm investigando como esses padrões são processados e aprendidos pelos falantes em diferentes aspectos e faixas etárias.

A próxima seção apresentará a noção de Aprendizagem Estatística e Habilidade Linguística Computacional e como esses dois construtos estão relacionados com os estudos sobre DNAs.

2.2 APRENDIZAGEM ESTATÍSTICA E HABILIDADE COMPUTACIONAL NA APRENDIZAGEM DE DNAS

As línguas naturais possuem inúmeras regularidades envolvendo aspectos prosódicos, fonológicos, morfológicos, sintáticos, etc. No estudo da Psicolinguística, muitas pesquisas têm investigado como essas regularidades são percebidas e aprendidas pelos falantes. A percepção e aprendizagem de regularidades, sejam elas linguísticas ou não, é conhecida como aprendizagem estatística.

A aprendizagem estatística ou no inglês, *Statistical Learning*, seria a capacidade cognitiva de perceber e extrair regularidades estatísticas a partir de padrões de dados observados no ambiente. Essa habilidade, segundo Aslin et al. (1999), nos possibilitaria detectar e utilizar propriedades distribucionais para fazer inferências, inclusive em eventos futuros, sobre diversos aspectos não linguísticos, assim como linguísticos. Nos aspectos linguísticos, alguns exemplos são: probabilidade transicional na segmentação de palavras (Saffran et al., 1996), coocorrência de dois elementos adjacentes (Van der Bos et al, 2012; Siegelman e Frost, 2015) e inclusive, não adjacentes (Gómez, 2002), que é o escopo desta pesquisa. Portanto, a aprendizagem estatística é o resultado da aprendizagem guiada por essa capacidade de percepção de regularidades no contínuo da fala, tanto em bebês, quanto em crianças e adultos.

Saffran et al. (1996) conseguiram demonstrar que bebês aos 8 meses de idade foram capazes de perceber coocorrências linguísticas apenas baseando-se em informações estatísticas, sem nenhuma outra pista fonológica adicional. Através dessas informações, os bebês puderam delimitar fronteiras de palavras contando apenas com a noção de probabilidade transicional (em inglês *Transitional Probability* ou *TP*), que fornece probabilidades estatísticas de uma sílaba (ou som) surgir após a outra, de acordo com a alta ou baixa coocorrência no contínuo de fala. Saffran et al. (1996), trazem como exemplo o sintagma nominal *pretty baby*, onde, estatisticamente, a sílaba *ty* aparece com muito mais frequência logo após de *pre*, do que *ba* surgindo logo após *ty*. Ou seja, se há uma alta probabilidade de uma sílaba aparecer após determinada sílaba, há uma grande possibilidade delas fazerem parte de uma mesma palavra. No entanto, se essa sequência for de baixa probabilidade, provavelmente essas sílabas não

pertencem à mesma palavra. Com esses resultados, os autores defendem que a aprendizagem estatística parece ser um mecanismo de natureza inata que interage com a experiência do indivíduo, a fim de promover um desenvolvimento eficiente durante a aquisição de linguagem.

Ainda sobre mecanismos de natureza inata, Endress et al. (2009) defendem que dispomos de mecanismos perceptuais, de domínios mais gerais e outros mais específicos, que viabilizam a percepção do falante a uma regularidade e sua probabilidade de ocorrência, de forma que ela possa ser aprendida e adquirida mais tarde. Assim, as regularidades presentes no ambiente são percebidas a partir de mecanismos de domínio geral (de atenção, por exemplo), passando por mecanismos que possam ser utilizados por dois outros mecanismos de domínio especializado, como por exemplo, reconhecer repetições em uma sequência de elementos, podendo essas ser de natureza visual ou auditiva. Por outro lado, um mecanismo específico seria recrutado quando outros domínios não são capazes de processar a informação. Imagine uma sequência em que há uma repetição de cor. A percepção da regularidade envolve um mecanismo mais geral ou compartilhado com outros domínios, porém, o reconhecimento das cores só pode ser feito por um mecanismo específico visual.

Endress et al. (2009) caracterizam os mecanismos especializados como um tipo de mecanismo que possa ser específico ou geral, mas que atende a apenas uma função. Como exemplo no trabalho em questão, os pesquisadores apresentam dois exemplos que parecem ser bastante pertinentes para a aprendizagem estatística de regularidades: o primeiro é o que percebe relações de coocorrências ou dependências (adjacentes ou não adjacentes) e consegue aplicar essas relações em novos ambientes, ou seja, com itens inéditos (como veremos ainda nesta seção sobre o estudo de Marcus et al., (1999)); e o segundo que é capaz de reconhecer a posição dos elementos dispostos em uma regularidade, o que os autores defendem estar sempre em posição de fronteiras (como veremos mais adiante em outros estudos na seção 2.3). Os autores consideram, então, os mecanismos especializados e os mecanismos que são compartilhados por domínios distintos como mecanismos perceptuais e de memória primitiva (em inglês, *Perceptual or Memory Primitive*), também conhecidos como POMPs.

Segundo os autores, os POMPs são considerados primitivos porque esses parecem ser de natureza inata, além de estarem presentes em espécies de animais não-humanos. Dessa forma, os autores defendem que, apesar desses mecanismos não serem necessariamente de natureza linguística, os humanos recrutam alguns desses POMPs para a faculdade da linguagem no processo de aquisição para computarem informações que serão organizadas e abstraídas resultando na aprendizagem de regularidades gramaticais ao longo da vida do falante. Portanto, mecanismos de domínio geral - que possibilitam a percepção de diversas regularidades não

linguísticas e linguísticas, visuais ou auditivas, adjacentes ou não adjacentes parecem ser essenciais para a aprendizagem estatística, e, conseqüentemente, para a aquisição de linguagem.

Frost e colaboradores (2015) e Siegelman e Frost (2015) também sugerem que a aprendizagem estatística não faz parte de um mecanismo unitário, mas que, na verdade, as computações necessárias para esse tipo de aprendizagem serão especificadas de acordo com a natureza da modalidade e tipo de estímulo. Por exemplo, Frost et al. (2015) defendem que, ao percebermos e computarmos estímulos de natureza auditiva, uma sensibilidade temporal deve ser maior do que a sensibilidade espacial, o que seria o contrário, caso a natureza dos estímulos fosse visual. A diferença entre estímulos linguísticos ou não, dessa forma, também receberia tratamento diferente, recrutando redes neurais separadas, a fim de processar essas regularidades a partir de sua especificidade, resultando em sua aprendizagem.

Siegelman e Frost (2015) foram um pouco além quando, através de atividades distintas para medir aprendizagem estatística², propuseram que o sucesso do desempenho em dada tarefa por um participante não poderia prever o mesmo desempenho em outra tarefa que utilizou modalidade ou estímulos distintos pelo mesmo participante. Dessa forma, os autores sugerem que, apesar de possuímos uma habilidade computacional de aprendizagem estatística, esse mecanismo parece ser mais composicional do que unitário. Em outras palavras, podemos assumir que temos a capacidade de perceber e extrair as regularidades à nossa volta, porém os caminhos percorridos entre essa percepção e aprendizagem podem ser distintos a depender do tipo de informação a que estamos sendo expostos.

Os autores também defendem, a partir desse mesmo estudo, que a aprendizagem estatística não parece estar embutida (*nested*, em inglês) em nenhuma outra habilidade cognitiva geral, como por exemplo, inteligência e memória de trabalho. Essa conclusão partiu da ausência de correlação entre os resultados das diferentes tarefas aplicadas no estudo juntamente com teste de avaliação de habilidades cognitivas gerais³. Thiessen et al. (2016) não condicionam a aprendizagem estatística a essas habilidades, porém, eles reconhecem que o aprimoramento dessas habilidades compromete a sensibilidade desse tipo de aprendizagem.

² Nesse trabalho, os pesquisadores cruzaram as seguintes tarefas com as respectivas modalidades: auditório - não verbal-adjacente, auditório - verbal - adjacente, auditório - verbal - não adjacente, visual - não verbal - adjacente. Não foi encontrada nenhuma correlação de desempenho.

³ Os testes aplicados em Siegelman e Frost (2015) foram: *Raven's Advanced Progressive Matrices* (para inteligência), *Digit Span* (para memória de trabalho), *Verbal Working Memory Task (VWM)* (um tipo de teste para detectar deficiência intelectual), *Rapid Automatized Naming (RAN)* (um teste adicional para avaliar deficiência intelectual), *Syntactic Processing Task* (para avaliar deficiência de aprendizagem) e *Switch Task* (para avaliar funções executivas, como controle inibitório, por exemplo).

Endress e colegas (2009) também assumem que é necessário mais do que a memória de trabalho para dar conta do processamento das regularidades, e isso eles também atribuem ao POMP de memória primitiva. Nessa lógica, eles defendem que se contássemos somente com a memória de trabalho para esse tipo de aprendizagem, bastaria termos a habilidade de armazenar certas ocorrências para aprendermos qualquer tipo de regularidade. No entanto, temos evidências de que isso não acontece, como veremos na seção 2.3. Uddén et al. (2012) também defendem que dispomos de uma memória especializada para o processamento e computação de dependências não adjacentes, que pode ser recrutada quando essas possuem um longo material interveniente ou são partes de uma estrutura aninhada ou cruzada⁴. No entanto, os próprios autores enfatizam que essa questão requer mais estudos empíricos.

Quando nos deparamos, então, com padrões linguísticos não adjacentes, espera-se que a aprendizagem estatística tenha o(s) mecanismo(s) recrutado(s) para a percepção e extração dessas regularidades. Embora veremos, na seção 2.3, que a aprendizagem de dependências não adjacentes possui algumas particularidades em relação às adjacentes, a viabilidade para sua aprendizagem conta com as habilidades estatísticas mencionadas acima. Estudos sugerem que possuímos a capacidade de perceber regularidades de DNAs já aos 11 meses de vida (Name et al., 2011; Laguardia et al., 2013; Name e Shi, 2015; Laguardia et al., 2015).

Neste presente trabalho, portanto, assumimos a aprendizagem estatística como um mecanismo de papel fundamental na aprendizagem de DNAs.

Nessa linha de raciocínio, assumimos que, além de um mecanismo perceptual, possuímos uma habilidade linguística computacional que possibilita o tratamento dos padrões ou regularidades percebidos como uma abstração e generalização de regras a serem empregadas mais tarde em outros contextos com novos exemplares (Marcus et al., 1999, Peña et al., 2002; Teixeira, 2012). Dessa forma, defendemos que dispomos de mecanismo(s) perceptual(ais) para aprendizagem estatística de regularidades não-linguísticas, e que, quando esses estímulos forem de natureza linguística, além desse(s) mecanismo(s), eles serão tratados por mecanismos específicos linguísticos e computacionais que serão aprimorados ao longo do desenvolvimento cognitivo e linguístico do falante (Van Der Kant et al., 2020).

Sobre esse tratamento diferenciado em relação a aspectos linguísticos, temos Marcus et al. (1999) que investigaram para bebês de 7 meses e como eles se comportavam em relação a padrões de sequências ABA ou ABB para analisar a percepção e abstração de regras em relação à segmentação de palavras. Para isso, foram conduzidos três experimentos. No primeiro experimento, os bebês foram submetidos à técnica de fixação visual, após a fase de

⁴ Veremos mais sobre essas estruturas na seção 2.3, subseção 2.3.4

familiarização, com dois pares de três palavras monossílabas de uma gramática artificial. As sequências da fase teste eram compostas por palavras inéditas, porém, que seguiam o mesmo padrão fonético das palavras da fase de familiarização (ex. [+consonantal, + vozeada] [+consonantal,- vozeada] [+consonantal, + vozeada]). Dos 16 participantes, 15 fixaram por mais tempo em sequências inconsistentes com as da fase de familiarização. Para eliminar a influência dos aspectos fonéticos e investigar se a aprendizagem estatística daria conta de extrair as regularidades das sequências sem essas pistas, o segundo experimento contou com sequências na fase teste que não eram compatíveis foneticamente com as da fase de familiarização (como no Experimento 1), embora elas ainda atendessem aos padrões ABA e ABB. Os resultados obtidos foram os mesmos do Experimento 1, sugerindo que os bebês aos sete meses de idade teriam capacidade de extrair regularidades de sequências de palavras mesmo na ausência de regularidades fonéticas.

Assim, nos dois primeiros experimentos desse estudo, os autores concluíram que, para a percepção dessas regularidades, a aprendizagem estatística seria um dos mecanismos envolvidos. Porém, para a generalização das regras subjacentes nas quais essas regularidades estariam inseridas, seria necessário um outro tipo de mecanismo. Para essa investigação, os autores criaram duas sequências com duplicação, ABB e AAB. Dessa forma, se os bebês levassem em conta somente a informação de duplicação de dois elementos nos dois primeiros experimentos (sequências ABA e ABB), ou seja, uma repetição, esses novos dois padrões utilizados (ABB e AAB) não poderiam ser diferenciados um do outro. Além disso, os autores defendem que, por utilizarem palavras inéditas na fase teste, os bebês também não poderiam contar com a probabilidade transicional entre os itens adjacentes, já que era a primeira vez que seriam expostos a essas sequências. Como resultado, todos os bebês participantes olharam por mais tempo para as sequências com discrepâncias em relação à sequência da fase de familiarização. Para os pesquisadores, esse resultado não questiona a importância da aprendizagem estatística para percepção e extração de regularidades nas línguas. No entanto, a presença de um mecanismo com uma computação algébrica que trate os estímulos de forma que eles estejam inseridos em uma regra abstrata parece ser plausível. Ou seja, Marcus et al. (1999) defendem que os bebês precisariam de, pelo menos, dois tipos de mecanismos de aprendizagem disponíveis, um para rastrear regularidades estatísticas e outro para manipular as variáveis envolvidas nessas regularidades.

Na próxima seção, veremos com mais detalhes as descobertas sobre a aprendizagem de DNAs e como os conceitos de aprendizagem estatística e habilidade computacional ainda geram discussões em diferentes estudos.

2.3 A APRENDIZAGEM DE DNAs

Muitas pesquisas dedicadas à investigação de aprendizagem de DNAs se voltam principalmente para bebês - enquanto muitos processos de aquisição de linguagem estão em desenvolvimento - e adultos, quando pelo menos a primeira língua já está consolidada. A seguir, veremos alguns desses estudos e como suas descobertas contribuíram com a literatura sobre a aquisição e aprendizagem de DNAs nessas faixas etárias⁵.

Em estudos com DNAs e categorização de palavras com bebês de 11 meses (Name et al., 2011; Name e Shi, 2015), as pesquisadoras atribuíram o bom desempenho dos bebês nas tarefas a ambas as habilidades, estatística e algébrica ou computacional. Em uma tarefa de fixação visual, os bebês foram expostos às DNAs probabilísticas em um DP que continha dois determinantes para uma vogal temática distinta, e, ainda assim, puderam atribuir itens intervenientes inéditos a suas respectivas DNAs. Mais especificamente, em Name et al. (2011), trinta bebês canadenses de 11 meses participaram do experimento e foram familiarizados com duas DNAs probabilísticas cada. Os participantes foram divididos em dois grupos e cada grupo contou com duas DNAs formadas por dois determinantes do português brasileiro e uma terminação vocálica. (grupo 1 - *nosso/seu X-e* e *esse/meu X-a*; grupo 2 - *nosso/seu X-a* e *esse/meu X-e*). Na fase teste, onde foi aplicada a tarefa de fixação visual, foram apresentados aos bebês, seis *trials* gramaticais (congruentes aos estímulos da fase de familiarização) e seis agramaticais (incongruentes aos estímulos da fase de familiarização). As DNAs agramaticais eram formadas pelas DNAs as quais o grupo oposto foi exposto. Assim, se o participante tivesse sido familiarizado com a DNA *nosso Xe*, como sequência agramatical, eles ouviam *nosso guika*. Os elementos intervenientes da fase teste foram todos inéditos. Como resultado, ambos os grupos apresentaram tempo de fixação para as DNAs agramaticais estatisticamente maior do que para as gramaticais. As pesquisadoras interpretaram que, o fato dos elementos intervenientes terem sido inéditos durante a fase teste pode indicar que os bebês rastrearam as DNAs e usaram essa informação para categorizar novas palavras em uma classe gramatical apropriada (no caso do trabalho, um nome).

Peña e colaboradores (2002), em consonância com Marcus et al. (1999), também defenderam que, além da habilidade de aprendizagem estatística, seria necessário uma habilidade computacional para o tratamento dos estímulos linguísticos a fim de generalizar

⁵ Abordaremos a idade de 7 anos, que é a faixa etária deste trabalho, mais especificamente nos próximos capítulos.

regras subjacentes às regularidades (no caso, DNAs). Nesse estudo, que contou com cinco experimentos com adultos, os pesquisadores olharam para a capacidade dos participantes de segmentar palavras em um contínuo de fala (Saffran et al. 1996), onde havia uma DNA presente. No primeiro experimento, os participantes tiveram bom desempenho ao escolherem as palavras que continham a DNA. Já no segundo experimento, para testar se os participantes eram capazes de generalizar a relação de dependência entre os elementos, os pesquisadores adicionaram novos itens intervenientes durante o teste para as sequências com as DNAs e partes de palavras com itens conhecidos que tinham sido usados durante a familiarização. Como resultado, os participantes falharam em escolher as palavras que continham as regras da DNA. Assim, no experimento 3, os pesquisadores adicionaram pausas de 25ms entre as palavras, que por mais que fossem imperceptíveis para os participantes, poderiam funcionar como pistas adicionais para a DNA. Os resultados mostraram que as pausas inseridas entre as palavras testadas foram facilitadoras para a generalização de regras. Para testar então, se pausas seriam mais pertinentes para a generalização da regra do que o tempo de exposição aos estímulos (ou seja, as regularidades), os pesquisadores, nos experimentos 4 e 5, respectivamente, triplicaram⁶ o tempo de exposição dos participantes aos estímulos durante a familiarização (no entanto, sem pausas) e diminuíram significativamente⁷ o tempo de exposição com a adição de pausas entre as palavras durante a familiarização. Como resultado, os participantes do experimento 4 falharam em generalizar a regra contida na DNA, enquanto os participantes do experimento 5 obtiveram sucesso.

2.3.1 Pistas fonológicas: pausas e fronteiras como facilitadores

A esses resultados dos experimentos 4 e 5, mencionados acima, Peña e colaboradores (2002) atribuíram uma computação algébrica que seria capaz de computar os elementos linguísticos contidos em regra, juntamente com a pausa entre as palavras, a fim de facilitar esse processo. Vale lembrar que, como vimos na seção 2.2, Endress et al. (2009) defendem que, por meio de um mecanismo especializado em reconhecer itens em fronteiras, somos capazes de fazer esse tipo de computação. Contudo, eles não atribuem esse mecanismo exclusivamente a elementos linguísticos.

Há um corpo de pesquisas, no entanto, que defende que a aprendizagem estatística pode ser suficiente para a generalização de regras quando na presença de itens inéditos, seja em

⁶ De 10 min para 30 min.

⁷ De 10 min para 2 min.

processos de segmentação de palavra, categorização gramatical ou processamento sintático (Onnis et al., 2004; Onnis et al., 2005; Frost e Monaghan, 2016; Frost et al., 2020), embora eles reconheçam que alguns desses processos possam ser mais complexos que outros. Onnis et al. (2004) defendem que a aprendizagem estatística extrai informações que sejam mais salientes que outras, quando na presença de mais de uma regularidade. Por exemplo, pistas fonológicas envolvidas nos itens de uma co-dependência podem ser mais eficazes do que uma relação de DNA no processamento de segmentação de palavras. Assim, Onnis et al. (2005) discutiram os resultados de Peña et al. (2002), argumentando que a relação de dependência não-adjacente não seria o fator determinante para o sucesso da tarefa de segmentação de palavras, já que os participantes poderiam ter retirado informações do contínuo de fala de um ambiente em que plosivas fossem mais presentes em posições iniciais (/p/, por exemplo). Ou seja, apenas a informação fonética de posição inicial de palavra pode ter sido suficiente para um bom resultado na atividade, ao passo que a aprendizagem da dependência não-adjacente não tenha tido um papel preponderante para o bom desempenho dos participantes.

Wang e Mintz (2018) demonstraram que ambas dependências adjacentes (doravante, DAs) e não adjacentes foram aprendidas quando apresentadas isoladamente. Embora eles alegam que o tipo de informação estatística contida em ambas DAs e DNAs possa ser diferente, para ambos os padrões localizados em fronteiras, são recrutados os mesmos mecanismos envolvidos na aprendizagem estatística.

Pausas parecem ser eficientes quando a tarefa é, portanto, segmentar palavras ou reconhecer uma DA. Porém, quando essas pausas estão inseridas entre os elementos co-dependentes de uma DNA, a sensibilidade para percepção desse padrão pode trazer uma dificuldade a mais. Bebês de 17 meses, mas não de 14 meses, conseguiram perceber uma DNA do francês contida em dois sintagmas fonológicos distintos, onde os padrões se relacionavam entre determinantes e auxiliares (Van Heugten et al., 2010). Novamente, a relação de dependência é estabelecida em fronteiras, mas por possuírem uma pausa entre essas fronteiras (sintagmas fonológicos), isso pode levar um tempo maior para o bebê conseguir estabelecer uma relação entre os dois elementos envolvidos na DNA.

Em um outro estudo, dessa vez com adultos, Van der Bos et al. (2012) associaram pistas fonológicas a DNAs probabilísticas, o que foi fundamental para a aprendizagem delas quando comparadas às DNAs determinísticas. Quando ambos os padrões de DAs e DNAs não dispunham de informações salientes ou adicionais, os participantes falharam em reconhecer as DNAs, tendo bom desempenho somente com as DAs. No entanto, em experimentos subsequentes, os participantes foram bem sucedidos ao reconhecerem as DNAs quando estas

foram atreladas a padrões visuais (cor, forma, movimento, respectivamente) e a pistas fonológicas, onde os elementos B (em uma configuração aXb) de um padrão de DNA probabilístico eram dissílabos e rimavam entre si.

A partir dos estudos citados acima, parece possível afirmar que fronteiras e/ou pausas possam ser informações salientes para a percepção, extração e abstração de regras de dependências não adjacentes.

2.3.2 Variabilidade do X

Gómez (2002) em um estudo com bebês de 18 meses e adultos propôs que a alta variabilidade do material interveniente presente na DNA seria um fator facilitador no processo de aprendizagem desses padrões. Esse estudo apontou que, para que um indivíduo perceba uma regularidade entre itens não adjacentes, é necessário que se descarte anteriormente qualquer regularidade adjacente. Dessa forma, a autora criou o estímulo linguístico de DNAs com alguns cenários onde o número de materiais intervenientes se diferenciava entre os grupos de participantes. Com os adultos, foram testadas quatro possibilidades de número de materiais intervenientes distintos (2, 6, 12 ou 24 elementos) que pudessem facilitar a aprendizagem das DNAs; com os bebês foram testadas 3 possibilidades (3, 12 e 24 elementos). Com os participantes adultos, a acurácia das respostas do grupo que foi exposto a 24 elementos intervenientes distintos foi de 90%, enquanto os participantes dos outros grupos não ultrapassaram a marca de 65,5%. Para os bebês, o tempo de escuta para os participantes que foram expostos aos 24 elementos intervenientes mostrou um efeito de novidade durante a fase teste, sugerindo a aprendizagem das DNAs. Os participantes dos grupos com 3 e 12 elementos não apresentaram diferença significativa entre eles. A autora, no entanto, não defende que exista um “número mágico” de material interveniente que resulte em uma aprendizagem de DNAs. O que ela defende é que quanto menor a possibilidade do sujeito estar diante de uma DA (ou seja, uma maior variabilidade no material interveniente diminui a TP), maior a previsibilidade de que a relação de regularidade esteja em uma coocorrência de itens não adjacentes, causando, portanto, um redirecionamento de foco do participante.

Em outras palavras, a aprendizagem da DNA pode ocorrer, quando o falante descarta a possibilidade da regularidade em questão não formar uma dependência adjacente. E a alta variabilidade do material interveniente (X) que fornece essa informação ao falante de que aX ou Xb não formam DAs.

2.3.3 A aprendizagem de DAs e DNAs

Alguns estudos se propuseram a investigar em que medida a aprendizagem de dependências adjacentes e DNAs utilizam os mesmos recursos e mecanismos (Gómez, 2002; Onnis et al., 2004; Uddén et al. 2012; Wang e Mintz, 2018). De maneira geral, os resultados apontam que ambos os padrões de dependências, adjacente e não adjacente, parecem utilizar os mesmos mecanismos perceptuais. No entanto, as DNAs são consideradas mais complexas para a aprendizagem, podendo utilizar mais recursos atencionais e de memória. Uddén et al. (2012), em um estudo com DAs e DNAs, sugeriram que a aprendizagem de ambos os padrões não difere qualitativamente, ou seja, ambos os padrões apresentam condições perceptuais de serem adquiridos. No entanto, quantitativamente, os pesquisadores defendem que as DNAs requerem um tempo maior de exposição aos aprendizes, devido às particularidades já mencionadas acima. Deocampo et al. (2019), em um estudo com adultos que investigou a aprendizagem concorrente de DAs e DNAs, em experimentos com estímulos visuais-espaciais e visuais-verbais (separadamente), encontraram aprendizagem para ambos os tipos de dependência. Em relação à modalidade dos estímulos - visuais-espaciais e visuais-verbais, não foram encontradas diferenças significativas quanto à aprendizagem das DAs e das DNAs. O que se observou como diferença, no entanto, foi que, apesar dos resultados apontarem para a aprendizagem de ambas as dependências, as DAs parecem ter sido mais consolidadas do que as DNAs pelos participantes, quando analisados os julgamentos de gramaticalidade dos dois padrões. Para as DNAs, os participantes parecem ter sido mais flexíveis em julgar a gramaticalidade, ou seja, eles julgaram mais sequências de DNAs agramaticais como gramaticais do que as DAs (experimento 2). Os autores atribuem a essa diferença algumas suposições: (a) em consonância com Gómez (2002), as DAs formam representações mais precocemente do que as DNAs; (b) as DAs e DNAs podem, por vezes, ter processos parcialmente dissociados, tendo as DNAs um grau de complexidade maior; e (c) as representações formadas a partir dos padrões podem mudar ao longo do tempo de exposição, com uma tendência maior de declínio para as DNAs, tornando-se necessário um reforço mais contínuo.

2.3.4 Condições favoráveis para a aprendizagem de DNAs

Uma das questões dos estudos que investigam aprendizagem estatística e aprendizagem de DNAs é como essas regularidades que são separadas por um material interveniente são

computadas para a sua aquisição, e quais são as condições propícias e necessárias nas línguas para que isso ocorra. Newport et al. (2004), em um estudo abrangente com bebês e adultos, perceberam que a capacidade dos participantes em aprender DNAs dependia não do tempo de familiarização, mas de aspectos específicos de regularidades encontrados nesses padrões. Por exemplo, eles defenderam que a aprendizagem aconteceu quando as DNAs mantinham relações consonantais ao passo que o material interveniente era vocálico, ou em uma configuração oposta, tendo os elementos *a* e *b* vocálicos e o X consonantal. Dessa forma, a percepção dos itens pertencentes a uma DNA poderiam estar aplicada nos princípios da Gestalt⁸, onde elementos dependentes são perceptualmente similares. Isso levou os pesquisadores a concluir que as línguas possuem regularidades em padrões de DNAs que os indivíduos são capazes de processar. Assim, as relações fonológicas nesses padrões não seriam arbitrárias nas línguas.

De Vries et al. (2012) também defenderam que as línguas naturais acompanham a nossa capacidade de processamento de DNAs. Eles pesquisaram DNAs aninhadas e cruzadas⁹. Segundo esse estudo, quando as sequências de aninhadas continham três DNAs, os participantes não obtiveram sucesso acima da chance em recobrar o item B da segunda dependência (ex. em negrito $A_1A_2A_3B_3\mathbf{B}_2B_1$). Isso porque, a dependência nesta posição se assemelha ao efeito da ausência de verbo¹⁰ (em inglês, *the missing-verb effect*), que nas línguas naturais se dá quando há muitas dependências não resolvidas e há uma sobrecarga de memória no processamento também devido à distância entre elas. Não coincidentemente, o número de ocorrências de DNAs aninhadas nas línguas naturais normalmente não ultrapassa de duas.

Em suma, acerca das descobertas e direcionamentos sobre aprendizagem de DNAs na literatura, podemos dizer que: (i) a habilidade de perceber e extrair regularidades de padrões no contínuo da fala se deveria à aprendizagem estatística; (ii) sobre a abstração e generalização de regras, ainda existe o debate se a aprendizagem estatística também seria responsável por isso ou seria necessária uma computação algébrica, de especialidade linguística; (iii) pistas prosódicas

⁸ Proposta pelo psicólogo Rudolf Arnheim (1954), é uma teoria também conhecida como a “Lei da Simplicidade” ou “o todo unificado” e conta com 6 princípios perceptuais: similaridade (agrupamento), continuidade (sequência), fechamento (previsão do todo mesmo que não completo), proximidade (unidade criada a partir de elementos), figura (representação de imagem a partir de itens dispostos de determinada maneira) e simetria (busca por equilíbrio).

⁹ DNAs aninhadas são caracterizadas quando o material interveniente é uma outra estrutura sintática, o que acaba deixando os elementos da dependência mais distantes. Ex: O gato, [que a minha mãe resgatou e levou para casa], fugiu de novo (configuração $A_1A_2B_2B_1$). DNAs cruzadas são mais raras em línguas naturais mas são caracterizadas quando as dependências intercalam seus elementos em uma configuração $A_1A_2B_1B_2$.

¹⁰ Exemplo de de Vries et al. (2012): “The apartment that₁ the maid who₂ the service₃ had sent over₃ _₂ was well decorated₁” que é agramatical, não parece se diferenciar da mesma versão gramatical em termos de entendimento por falantes do inglês “The apartment that₁ the maid who₂ the service₃ had sent over₃ was cleaning every week₂ was well decorated₁”.

e fonológicas como pausas, padrões fonotáticos, regularidades fonéticas entre os itens da dependência seriam facilitadores para a percepção e aprendizagem de DNAs; (iv) a alta variabilidade do material interveniente (X) seria um facilitador para a aprendizagem de DNAs, principalmente, por descartar para o falante, a possibilidade dele estar diante de uma DA; (v) muitos pesquisadores defendem que ambas DAs e DNAs são processadas pelos mesmos mecanismos, embora seja amplamente reconhecido que DNAs demandam processos e recursos cognitivos mais complexos e; (vi) as línguas naturais atenderiam aos requisitos perceptuais e de processamento para que sejamos capazes de perceber e aprender DNAs presentes nas línguas naturais.

2.4 CONCLUSÃO

Neste capítulo, foram apresentadas as DNAs e algumas possibilidades de ocorrências em línguas naturais, a fim de demonstrar que, além de muito presente nas línguas, são regularidades bastante complexas.

Vimos como a aprendizagem estatística é atribuída na percepção e extração de regularidades presentes em ambientes linguísticos ou não, o que configuraria uma habilidade de domínio geral, além de mecanismos de domínios compartilhados e especializados. No entanto, vimos também que alguns estudos defendem que possuímos habilidades linguísticas especializadas que tratam essas regularidades como componentes de uma regra. Portanto, uma habilidade computacional ou algébrica seria necessária para abstrações e generalizações de regras que os componentes linguísticos envolvem.

Foi apresentado, de maneira geral, o que os estudos apontam como evidências sobre aprendizagem de DNAs em bebês e adultos.

No próximo capítulo, será apresentado um apanhado sobre o perfil cognitivo e metacognitivo, além da maturação cerebral da faixa etária de 7 anos, a fim de contextualizar os desafios que os estudos que investigam essa fase de desenvolvimento enfrentam.

Para analisarmos o que já foi investigado sobre a aprendizagem de DNAs por crianças de 7 anos, este trabalho conta com dois estudos de base, que são Lammertink et al. (2019) e Marimon et al. (2021). Ambos são considerados de base por terem examinado a aprendizagem de DNAs por participantes de faixa etária de 7 anos de idade, a mesma que aqui está sendo tratada. Esses trabalhos servirão de guia para discussões de como a aprendizagem de DNAs por crianças nessa idade foi investigada, e o que esses estudos deixaram em aberto. Como mencionado anteriormente, as questões suscitadas a partir desses dois trabalhos e que serão

apresentadas no próximo capítulo são os padrões determinísticos e probabilísticos de DNAs, o papel da atenção aos estímulos durante a aprendizagem implícita de crianças dessa faixa etária e a observação da aplicação da tarefa de Escolha Forçada como atividade experimental para esse tipo de investigação.

Ao longo das seções 4.2, 4.3 e 4.4, outros trabalhos serão apresentados em diferentes momentos, a fim de destacar o ponto de interesse específico para cada questão listada acima.

Primeiramente, será apresentado um breve contexto cognitivo sobre a idade investigada, visto que estudos apontam mudanças cognitivas e metacognitivas em relação ao cérebro de bebês, crianças mais novas e de adultos. Assim, entender os processos cognitivos e metacognitivos em desenvolvimento pode auxiliar na compreensão e análise dos resultados obtidos.

Em seguida, serão apresentados dois tipos de DNAs presentes em estudos empíricos: os padrões de DNAs determinísticos e os padrões de DNAs probabilísticos. A seção 3.2 busca, além de delinear as características de ambos os padrões, também discutir a implicação do uso de cada um deles. Dessa forma, mostra-se uma lacuna em estudos empíricos de padrões de DNAs probabilísticos na faixa etária de crianças em idade escolar.

Posteriormente, na seção 3.3, é apresentada uma revisão de literatura sobre o papel da atenção na aprendizagem de DNAs e como ela pode ser manipulada durante as tarefas experimentais. Essa seção contempla estudos em diferentes faixas etárias, e sinaliza a ausência de dados para a faixa etária dos 7 anos de idade, quando a atenção está relacionada à aprendizagem de DNAs.

E, por fim, a partir da apresentação de conceitos de aprendizagem implícita e intuitividade na seção 3.4, discute-se o papel das instruções nas tarefas experimentais, e quais as suas consequências. Além disso, busca-se compreender conceitos como tarefas *online* e *offline* e medidas diretas e indiretas, a fim de avaliar a aplicabilidade de uma tarefa de Escolha Forçada na investigação de aprendizagem implícita de DNAs por crianças em idade escolar.

3 O PERFIL COGNITIVO E METACOGNITIVO DE CRIANÇAS DE 7 ANOS

No capítulo anterior, foi apresentada uma revisão de literatura sobre aprendizagem de DNAs abordando diversos aspectos. Embora haja um corpo de pesquisa robusto acerca do tema, quando olhamos para as idades investigadas, nos deparamos mais com estudos com bebês e adultos. Dessa forma, crianças que já possuem uma idade mais avançada e que possuem, ao menos, uma língua consolidada não parecem ser frequentemente objeto de investigação. Em relação à aprendizagem de DNAs, pudemos contar com poucas referências (Lammertink et al., 2019; Paul et al., 2019, Schaadt et al., 2021; Marimon et al., 2021) que se voltaram para a faixa etária dos 7 anos e idades adjacentes. Dois desses estudos (Lammertink et al., 2019; Marimon et al., 2021), portanto, serão adotados neste trabalho como estudos de base. Esses estudos serão detalhadamente descritos no próximo capítulo, na seção 4.2, quando trataremos das questões levantadas nesta dissertação. Assim, este capítulo propõe-se a contextualizar o perfil cognitivo e metacognitivo de crianças de 7 anos de idade, através de estudos que investigaram aprendizagem em diferentes situações (categorização, segmentação de palavras, aprendizagem implícita, etc), a fim justificar a importância e a complexidade de examinar como se dá a aprendizagem nessa faixa etária, além de quais recursos estão disponíveis durante esse processo.

3.1 PROCESSOS COGNITIVOS E METACOGNITIVOS NA INFÂNCIA: O PAPEL DA IDADE NA APRENDIZAGEM

Na faixa etária dos 7 anos de idade, considera-se que a primeira língua já esteja consolidada, mas que ainda há muitos processos cognitivos (Schaadt et al., 2019) e metacognitivos (Bialystok, 1986; Edwards e Kirkpatrick, 1999) em desenvolvimento. A maturação do córtex pré-frontal traz grande impacto nesta idade, que já se assemelha mais com um cérebro adulto do que com um de bebê em relação à plasticidade sináptica (Huttenlocher, 1990) e controle inibitório (Thiessen et al., 2016), por exemplo.

Sobre essa maturação cerebral, McNealy e colaboradores (2011) realizaram um estudo sobre segmentação de palavras que contemplou crianças de 5 a 7 anos (média de idade de 6a10m), 9 e 10 anos (média de idade 10a), 12 e 13 anos (média de idade 12a11m) e adultos. Neste estudo, eles realizaram juntamente com uma tarefa experimental comportamental de discriminação de palavras, a análise de imagens de ressonância magnética funcional (fMRI) enquanto os participantes eram expostos aos estímulos linguísticos durante a fase de

familiarização. O objetivo era observar se as atividades cerebrais nesse tipo de tarefa - de segmentação de palavra - sofreriam alguma alteração com o avanço da idade. Além disso, eles criaram três condições, às quais os participantes de todos os grupos foram expostos: (1) condição de língua sem saliência (U); (2) condição de língua com saliência (S) e; (3) condição de sílabas aleatórias (R). A condição de língua sem saliência consistia em sequências de palavras formadas por três sílabas (estímulos de Saffran et al., 1996). Neste estímulo linguístico contínuo, os participantes só poderiam segmentar as palavras através da probabilidade transicional entre as sílabas. Importante ressaltar que esse estudo só contou com dependências adjacentes. Na condição de língua com saliência, os pesquisadores adicionaram *stress* na primeira sílaba da sequência de três sílabas em um terço das vezes que a mesma palavra ocorria, a fim de dar um pista prosódica aos participantes, visto que todos eram falantes de inglês e o padrão tônico desta língua tem em sua maioria palavras proparoxítonas. Na terceira condição, a de sílabas aleatórias, as palavras eram formadas somente por duas sílabas e a probabilidade transicional entre elas decrescia drasticamente (nas condições U e S, a probabilidade era 1 e na R, no máximo, 0,22). A ordem das condições foi contrabalanceada entre os participantes e nenhuma condição compartilhava da mesma palavra que a outra.

Esse estudo trouxe descobertas interessantes para a idade aqui investigada. A primeira delas, em relação à atividade cerebral, revelou que as crianças na faixa etária dos 6 anos usaram mais regiões cerebrais do hemisfério direito (giro temporal direito, ínsula direita, giro pré-central direito) do que todos os outros grupos de participantes. Enquanto os participantes de 10 anos ainda revelaram uma bilateralidade, os de 13 e os adultos ativaram majoritariamente o lado esquerdo do cérebro. Essa descoberta corrobora com a ideia de que nessa idade, ainda há transformações neurais, cognitivas e sinápticas em relação à aprendizagem. Dessa forma, é reforçada a ideia de que crianças de 7 anos de idade pareçam estar em um estágio cerebral maturacional intermediário entre bebês e adultos.

Uma outra descoberta interessante em relação à atividade cerebral no estudo de McNealy et al. (2011) foi que, diferentemente de todos os outros grupos (com uma diferença menor entre os participantes de 10 anos), os participantes de 6 anos revelaram uma sensibilidade aumentada em relação à condição R. Ou seja, mesmo que estatisticamente não houvesse relação forte de dependência entre as sílabas nessa condição, o cérebro parecia estar sensível e de prontidão caso aquelas sequências pudessem apresentar uma regularidade relevante e passível de aprendizagem. Isso sugere que crianças mais novas apresentam uma sensibilidade maior do que as mais velhas e adultos em relação ao ambiente, para extrair dele informações que possam ser usadas mais tarde, ainda que de maneira bastante implícita. Sobre

essa sensibilidade mais aguçada, Thiessen e colaboradores (2016) explicam que essa capacidade, também presente nos bebês, de codificar informações com mais ruídos, permite que os indivíduos considerem características idiossincráticas dos elementos, mas que de alguma forma não são comuns entre si. Ou, mesmo que haja relação comum entre os elementos, essa sensibilidade permite codificar outras informações dos estímulos que possam tornar-se cruciais em um futuro contexto. Thiessen et al. (2016) ainda afirmam que, apesar dessa capacidade resultar em um tempo necessário maior para a aprendizagem, ela também impede generalizações entre itens que não necessariamente compartilham relações e, permite generalizações entre itens que *a priori* não pareçam estar relacionados.

Por último, nesse mesmo estudo de McNealy et al. (2011), a tarefa comportamental de discriminação de palavras, que também pode ser entendida como uma tarefa de julgamento de gramaticalidade, os participantes deveriam responder “sim” ou “não” (se elas pareciam com palavras a que tinham sido expostos na fase anterior) quando escutavam oitos palavras de três sílabas congruentes com o período de familiarização e oito palavras formadas por sílabas finais de uma palavra com iniciais de uma outra palavra (em inglês são descritas como *partwords*). A ordem das palavras era aleatória. Os resultados dessa tarefa ficaram no nível da chance, podendo sugerir que, mesmo o cérebro mostrando ativações no momento da aprendizagem, a tarefa comportamental pode exigir algo mais entre um processo e outro que ainda não é possível constatar. No caso desse estudo, poderia ser uma análise metalinguística, por exemplo. Na seção 4.3, discutiremos alguns problemas pontuais, os quais tarefas comportamentais podem gerar a partir das instruções dadas aos participantes e que, por sua vez, possam comprometer resultados na investigação de aprendizagem implícita. Em outras palavras, o modo de como o teste comportamental é aplicado e como e quando as instruções são dadas, podem ajudar ou dificultar o bom desempenho do participante na tarefa.

Alguns estudos sugerem que os mecanismos de aprendizagem passam por modificações de procedimentos associativos até próximo aos 4 anos de idade para processos de aprendizagem mais controlados, ou baseados em regras, até os 7 anos de idade (Skeide e Friederici, 2016; Schaadt et al., 2019; Van Der Kant et al., 2020, Paul et al., 2020). Esses processos de aprendizagem baseados em regras, por sua vez, continuam em desenvolvimento ao longo dos anos.

Segundo Rabi e Minda (2014), aprendizagem associativa, em inglês *Associative Learning*, seria um processo de aprendizagem implícita, quando não é possível verbalizar as regras subjacentes, mas que permite o sujeito associar características a elementos, enquanto aprendizagem baseada em regras, em inglês *Rule-based Learning*, seria um processo, o qual o

sujeito tem acesso a uma regra que seja relativamente fácil de descrever verbalmente. Porém, é um processo que requer recursos cognitivos suficientes, assim como memória de trabalho e controle inibitório mais desenvolvidos.

Rabi e Minda (2014) descrevem que categorias formadas a partir de regras (*rule-based*) apresentam uma regra mais típica o possível (em inglês, o adjetivo usado é *optimal*), ou seja, é uma característica que se destaca facilmente entre outras a ponto de poder ser verbalizada, configurando uma regra que se diferencie dos outros itens que não são categorizados no mesmo conjunto ou classe. Por outro lado, itens, sejam eles linguísticos ou não, podem compartilhar características que possam ser categorizadas, mas que não necessariamente, verbalizadas. Por exemplo, itens que possuem uma característica comum a todos (objetos amarelos), mas que outras características não parecem ser compartilhadas entre si (objetos pequenos amarelos, objetos redondos amarelos, instrumentos musicais amarelos). Nesse caso, alguns pesquisadores argumentam que a categorização ocorre de forma implícita por meio de um sistema suficientemente sensível para identificar características comuns entre os elementos e agrupá-los em um mesmo conjunto, ainda que o falante não consiga verbalizar a razão pela qual esses itens compartilham uma característica regida por uma regra.

Roark e Holt (2019) investigaram como crianças entre 5 e 7 anos e adultos categorizavam sons baseando-se em sistemas integrativos de aprendizagem¹¹: o implícito e o explícito (controlado e baseado em regras). O objetivo era observar se crianças nessa faixa etária, que estão passando por maturações no córtex pré-frontal, integrariam os dois sistemas na tarefa de categorização. Além disso, a taxa de acurácia entre os dois grupos seria avaliada. Como resultado, os participantes adultos tiveram melhor desempenho em relação às crianças, que ficaram no nível da chance. Quando aplicados modelos de integração de informação para avaliar as estratégias dos participantes, os pesquisadores observaram que os adultos recorreram aos dois sistemas de aprendizagem, o implícito e o explícito, à medida que um fosse mais pertinente que o outro. As crianças, por outro lado, especialmente as mais velhas e que ficaram na chance ou abaixo da chance, ao invés de integrarem os dois sistemas, elas mudaram um pelo outro por diversas vezes durante a tarefa em uma tentativa de conseguir categorizar os

¹¹ COVIS (Competition between Verbal and Implicit Systems). Segundo esse modelo teórico (Paul e Maddox, 2011) de aprendizagem de categorização, o sistema verbal, ao aprender uma nova categoria, busca primeiramente por uma regra explícita para classificá-la. No entanto, se a regra subjacente não for facilmente percebida, o sistema implícito assume o controle, baseando-se na experiência. O sistema explícito, no entanto, não parece ser recrutado antes dos três anos de idade, e sua eficiência acompanha as transformações e desenvolvimento do córtex pré-frontal durante a infância e adolescência.

estímulos, o que pode ter comprometido a acurácia, visto que esse processo pode levar mais tempo.

Rabi e Minda (2014) também aplicaram uma tarefa de categorização em crianças de 4 a 11 anos, além de adultos. Para a faixa etária dos 7 anos, os participantes ficaram abaixo do nível da chance, diferentemente das crianças de 10 anos, que apresentaram resultados similares aos adultos. Além da maturação cerebral do córtex pré-frontal, os pesquisadores sugeriram que a diferença no desempenho na tarefa pode ter sofrido influência da memória de trabalho, que aos 7 anos e anterior a essa idade, ainda está em desenvolvimento.

Thiessen et al. (2016), em um artigo de revisão sobre aprendizagem e período crítico contra-argumentam a ideia de que os mecanismos de aprendizagem se alteram a partir de uma certa idade. Em contrapartida, eles alegam que os mecanismos envolvidos na aprendizagem acompanham os indivíduos ao longo da vida, porém, com uma perda de eficiência ao longo do tempo. À essa perda de eficiência, eles não atribuem somente o avanço da idade e processos cognitivos maturacionais, mas também questões mais individuais como experiência e recursos atencionais e de memória. Para exemplificar e sustentar essa hipótese, os autores sugerem que a aprendizagem estatística seria um desses mecanismos.

A relação entre experiência e aprendizagem estatística, segundo os autores, pode ser entendida como a ideia de que quanto menos linguisticamente experiente o indivíduo for, menos influências suas experiências prévias incidem sobre novos estímulos. Com isso, indivíduos menos experientes precisam dispor de um mecanismo de aprendizagem estatística mais sensível, para que o *input* seja analisado com menos vieses e, portanto, com menos chances de abstrair relações inexistentes e, consequentemente, gerar generalizações errôneas. Por outro lado, quanto mais experiência linguística o falante tiver, como por exemplo, crianças mais velhas e adultos, mais influência essa experiência exerce sobre novos estímulos. A experiência individual, portanto, implica parcialmente em uma compensação da perda de sensibilidade nos mecanismos de aprendizagem estatística (nem sempre de maneira tão eficiente) para percepção e extração de regularidades, a partir de conhecimentos já adquiridos e sedimentados.

Os recursos atencionais, segundo Thiessen et al. (2016), também vão se aprimorando com o avanço da idade. A fase adulta, por exemplo, conta com direcionamento de foco atencional mas desenvolvido do que crianças e bebês. No entanto, os autores ponderam que esse direcionamento de foco nem sempre é benéfico para a aprendizagem de uma nova língua. Isso porque os estímulos podem não corresponder às expectativas dos adultos, ao ponto de que ao focarem em características não relevantes dos estímulos, outras possam passar

despercebidas, e por sua vez, não serem extraídas. O controle inibitório plenamente desenvolvido também pode custar a alta eficiência da aprendizagem implícita, de maneira que os adultos, ao inibirem concorrentes que, em princípio, parecem carregar informações irrelevantes, acabam descartando outras possibilidades de relações e adaptações futuras. Crianças, por outro lado, por não inibirem esses competidores, acabam permitindo mais dessas ocorrências, fortalecendo o padrão dominante por ser continuamente reativado, promovendo a aprendizagem. Assim, os autores sugerem que o aumento dos controles atencional e inibitório pode resultar em uma perda de sensibilidade nos mecanismos de aprendizagem estatística ao longo do desenvolvimento cognitivo. Considerando que os recursos atencionais e inibitórios estão ligados também ao controle executivo, a maturação do córtex pré-frontal exerce bastante influência nessa mudança. E, levando em consideração que crianças de 7 anos de idade têm essa área cerebral ainda em desenvolvimento, é necessário buscar mais informações para se entender melhor em que medida a atenção interage com a aprendizagem estatística.

Quanto à aprendizagem implícita, no entanto, Reber e Allen (1993) assumem que ela ocorre desde muito cedo em nosso desenvolvimento e que esse tipo de aprendizagem não é relacionado à idade. Para investigar essa questão, Vinter e Perruchet (2000), conduziram experimentos de aprendizagem implícita com crianças de 4 a 10 anos e adultos em uma atividade de desenho, onde os participantes deveriam traçar formas geométricas a partir de padrões (ponto de partida e direção do traço), aos quais eles foram expostos durante a fase de treinamento. Os pesquisadores não observaram nenhuma diferença significativa entre os grupos de idades em relação ao tipo da tarefa. Dessa forma, eles observaram que não houve nenhuma interação entre capacidade em aprender um padrão (ponto de partida combinado com a direção da rotação do traço) e idade do participante, sugerindo que a capacidade de aprender implicitamente nos acompanha ao longo da vida. No entanto, eles não excluem a possibilidade de alguns tipos de aprendizagem implícita demandarem outros aspectos que sejam característicos a um desenvolvimento típico de determinada idade.

Acerca de processos metacognitivos que caracterizam a faixa etária dos 7 anos de idade, em um estudo seminal de Bialystok (1986) sobre consciência metalinguística em crianças de 5, 7 e 9 anos de idade, observou-se que as crianças de 7 anos se mostraram capazes de fazerem julgamentos gramaticais, mas não apresentaram bom desempenho em relação às crianças mais velhas de nove anos de idade em tarefas de correção. Esse estudo contou com duas fases de experimento. Na primeira fase, as crianças deveriam julgar a fala de um fantoche em dois níveis: se era gramatical (*grammatical* em inglês ou *G*) ou com sentido (*meaningful* em inglês ou *M*). Caso elas não fossem julgadas ou como gramatical ou com sentido, os

pesquisadores marcaram as sentenças como *g* (agramatical) ou *m* (sem sentido, em inglês *meaningless*). Dessa forma, era esperado que, uma sentença como “*If I am sick again tomorrow, I will have to see my fireman*” fosse julgada como (*Gm*), ou seja, gramatical mas sem sentido. Na segunda fase, que configurava a etapa de correção, as crianças eram instruídas a corrigirem algumas sentenças *gm* usadas na primeira fase do experimento, ou seja, agramaticais e sem sentido. Nessa segunda fase, quando foi necessário pensar sobre a regra e identificar o erro (de gramática e de sentido) na sentença, crianças de 7 anos mostraram mais dificuldade quando comparadas com o grupo de crianças de 9 anos. Isso sugere que, apesar de uma habilidade metalinguística já estar em desenvolvimento aos sete anos de idade, ela ainda não atingiu seu ápice.

Edwards & Kirkpatrick (1999) investigaram idades entre 4 e 12 anos, e encontram uma mudança drástica na habilidade metalinguística entre os 7, 8 e 9 anos de idade, sugerindo que nessa faixa etária pode existir um marco transicional em habilidades metacognitivas de análise e controle. Nesse experimento, as crianças eram instruídas a ouvirem uma história. Nessa história, 20 sentenças continham uma pseudopalavra no lugar de uma palavras do inglês (língua nativa dos participantes). Metade das palavras substituiu itens funcionais (preposições, conjunções, marcas de infinitivo) e a outra metade itens lexicais (verbos, adjetivos e nomes). As crianças deveriam pressionar um botão assim que ouvissem uma palavra “maluca” (pseudopalavra). Os resultados de acurácia dessa tarefa foram estatisticamente mais altos para os grupos que contemplavam crianças mais velhas (acima de 8 anos) quando comparados com as crianças mais novas (7 anos ou menos). No entanto, quando analisado se houve diferença de acurácia em relação às pseudopalavras que substituíram itens funcionais e lexicais, só a partir da faixa etária dos 9 anos de idade que houve mais acertos para substituições de itens funcionais, ou seja, as crianças mais velhas se mostraram mais sensíveis a violações morfosintáticas do que as mais novas, que julgaram os dois tipos de violações (itens funcionais e lexicais) de maneira similar.

Em um estudo com crianças brasileiras de 7 anos que buscou investigar a interação do processo de alfabetização com processos metalinguísticos, Barrera e Maluf (2003) conduziram tarefas que acessaram explicitamente a consciência fonológica, lexical e sintática dos participantes durante o início e fim do primeiro ano do ensino fundamental, quando normalmente as crianças estão no início da aprendizagem das habilidades de escrita e leitura. A seguir, apresentaremos brevemente os resultados relevantes para o nosso trabalho que as pesquisadoras encontraram.

As habilidades metalinguísticas avaliadas (fonológica, lexical e sintática) não foram correlacionadas, ou seja, um participante que teve um desempenho bom em uma das tarefas não necessariamente teve o mesmo desempenho nas outras. Uma das interpretações possíveis é que, processos metacognitivos independentes e distintos podem estar envolvidos de acordo com a tarefa em questão e a complexidade de análise. Entre as habilidades metalinguísticas testadas, a de consciência fonológica foi a que mais apresentou correlação com a habilidade de escrita e leitura. Tanto no início quanto no final do ano, bons indicadores de consciência fonológica interagiram com bons indicadores de habilidades de leitura e escrita. Em uma das tarefas para testar consciência metalinguística fonológica, os participantes precisavam selecionar, de um total de três palavras, duas que começavam com o mesmo som (aliteração) e duas que terminavam com o mesmo som (rima). As pesquisadoras observaram uma sensibilidade mais alta para as aliterações do que para as rimas, sugerindo que crianças não alfabetizadas ou em processo de alfabetização podem focar mais no som inicial das palavras do que no final. Em relação às rimas, foi observado que os participantes acertaram mais quando as palavras compartilhavam mais fonemas entre si do que apenas uma sílaba (ex. *torneira* e *cadeira* foram mais relacionadas do que *faca* e *boca*). Segundo as pesquisadoras, isso pode indicar que a quantidade de fonemas compartilhados pode ser um facilitador para percepção de rimas entre palavras para crianças nesse estágio de alfabetização. Acerca da consciência metalinguística lexical, as autoras aplicaram uma tarefa de segmentação de palavras de sentenças orais. Após a leitura dessas sentenças, elas eram apresentadas visualmente para as crianças sem segmentação das palavras para que os participantes colocassem um cartão colorido em cada palavra da sentença. Cada sentença continha de 3 a 5 palavras. Um resultado interessante foi que houve uma hipossegmentação entre os artigos ou conjunção e substantivos (ex. *e Pedro; um sapato*). Dessa forma, para crianças de 7 anos ainda no processo de alfabetização, o limite de palavra morfológica parece ainda não estar totalmente sedimentado, de forma que essa segmentação explícita pode ser feita a partir de uma noção de palavra fonológica ou prosódica. Portanto, crianças de 7 anos com desenvolvimento típico que já possuem uma língua consolidada e dispõem de recursos linguísticos com bastante eficiência para se expressarem, parecem precisar de processos mais explícitos (metacognitivos) para a aprendizagem de habilidades também explícitas, como a leitura e a escrita.

3.2 CONCLUSÃO

Diante de tudo isso, os processos metacognitivos em desenvolvimento combinados com processos educacionais formais de alfabetização parecem guiar as crianças dessa idade a pensar sobre a sua língua, em um momento que ainda sofrem transformações de processos cognitivos, neurais e de maturação cerebral. Para o presente estudo, os processos metacognitivos são secundários, visto que buscamos investigar a aprendizagem implícita resultante da aprendizagem estatística. Portanto, apesar de descrevermos diversas transformações pelas quais crianças de 7 anos passam nesse período, na análise de resultados só serão levados em conta aspectos que impactam os tipos de aprendizagem aqui investigados, ou seja, aprendizagem estatística e aprendizagem implícita.

No próximo capítulo, veremos de maneira mais aprofundada o papel da atenção na aprendizagem de DNAs e a investigação de aprendizagem implícita, especificamente na tarefa de Escolha Forçada, a fim de contribuir ainda mais com os dados apresentados até agora.

No entanto, iniciaremos a discussão com a diferença entre os padrões determinísticos e probabilísticos de DNAs em estudos empíricos, e como esses padrões são aplicados em estudos, visto que os estímulos linguísticos tratados neste presente trabalho contemplam o segundo tipo.

4 DNAs DETERMINÍSTICAS E PROBABILÍSTICAS, O PAPEL DA ATENÇÃO NA APRENDIZAGEM DE DNAs E A INVESTIGAÇÃO DE APRENDIZAGEM IMPLÍCITA

Como mencionado no capítulo anterior, a fim de analisarmos o que já foi investigado sobre a aprendizagem de DNAs por crianças de 7 anos, este trabalho adota dois estudos de base, que são Lammertink et al. (2019) e Marimon et al. (2021). Ambos são considerados de base por terem examinado a aprendizagem de DNAs por participantes de faixa etária de 7 anos de idade, a mesma que aqui está sendo tratada. Esses trabalhos serviram de guia para discussões de como a aprendizagem de DNAs por crianças nessa idade foi investigada, e o que esses estudos deixaram em aberto. Como mencionado anteriormente, as questões suscitadas a partir desses dois trabalhos e que serão apresentadas neste capítulo são: os padrões determinísticos e probabilísticos de DNAs, o papel da atenção aos estímulos durante a aprendizagem implícita de crianças dessa faixa etária e a avaliação da aplicação da tarefa de Escolha Forçada como atividade experimental para esse tipo de investigação.

Ao longo das seções 4.1, 4.2 e 4.3, outros trabalhos serão apresentados em diferentes momentos, a fim de destacar o ponto de interesse específico para cada questão listada acima.

Primeiramente, serão apresentados dois tipos de DNAs presentes em estudos empíricos: os padrões de DNAs determinísticos e os padrões de DNAs probabilísticos. A seção 4.1 busca, além de delinear as características de ambos os padrões, também discutir a implicação do uso de cada um deles. Dessa forma, mostra-se uma lacuna em estudos empíricos de padrões de DNAs probabilísticos na faixa etária de crianças em idade escolar.

Posteriormente, na seção 4.2, é apresentada uma revisão de literatura sobre o papel da atenção na aprendizagem de DNAs e como ela pode ser manipulada durante as tarefas experimentais. Essa seção contempla estudos em diferentes faixas etárias, e sinaliza a ausência de dados para a faixa etária dos 7 anos de idade, quando a atenção está relacionada à aprendizagem de DNAs.

E, por fim, a partir da apresentação de conceitos de aprendizagem implícita e intuitividade na seção 4.3, discute-se o papel das instruções nas tarefas experimentais, e quais as suas consequências. Além disso, busca-se compreender conceitos como tarefas *online* e *offline* e medidas diretas e indiretas, a fim de se avaliar a aplicabilidade de uma tarefa de Escolha Forçada na investigação de aprendizagem implícita de DNAs por crianças em idade escolar.

4.1 PADRÕES DETERMINÍSTICOS E PROBABILÍSTICOS DE DNAs

Como vimos no capítulo 2, na subseção 2.3.2, em um estudo seminal sobre a aprendizagem de DNAs, Gómez (2002) investigou qual seria a melhor condição de aprendizagem de DNAs por bebês e adultos em relação ao número de exemplares do material interveniente (o X da estrutura aXb) apresentados nos estímulos durante a fase de familiarização. Um dos resultados encontrados foi que quanto maior a variabilidade do material interveniente de uma DNA, mais facilidade teriam os participantes de aprenderem o padrão. Além de Gómez (2002), outros estudos (Gómez e Maye, 2005; Onnis et al., 2004; Newport e Saffran, 2004; López-Barroso et al., 2016; Grama e Wijnen, 2018; Lammertink et al., 2019) utilizaram estímulos linguísticos de gramática artificial para a observação da aprendizagem do padrão de DNAs de pseudopalavras, ainda que com objetivos distintos.

No entanto, o que todos esses e tantos outros estudos têm em comum é o fato de que eles investigaram padrões de DNAs determinísticos, ou seja, padrões em que um mesmo elemento a sempre se relaciona a um mesmo elemento b (em uma configuração aXb). Em outras palavras, a probabilidade de um elemento final surgir quando tal elemento inicial estiver presente é 1. Assim, não há possibilidade de variação de elementos que configurem um conjunto que sejam codependentes de outros elementos, que por sua vez também façam parte de um conjunto. Esse tipo de padrão de DNAs determinístico é muito usado na investigação de sensibilidade e aprendizagem de padrões por participantes por manter uma configuração mais superficial. Dessa forma, os pesquisadores conseguem isolar quaisquer influências e focar somente no resultado significativo dessa aprendizagem.

Por outro lado, as DNAs determinísticas deixam a desejar quando o objetivo é simular um ambiente de aprendizagem de padrão de DNA em uma língua natural, mesmo utilizando estímulos de uma língua artificial ou pseudonatural.

Como apresentado na seção 2.1, as DNAs mais comumente encontradas em línguas naturais são relações de codependência de classes ou categorias funcionais, estabelecendo, na maioria das vezes, relações de concordância morfossintática. E como o próprio nome já diz, classe ou categoria é um conjunto de exemplares que atendem a um mesmo requisito. Assim, em uma classe de determinantes no Português Brasileiro, por exemplo, encontramos diversos itens: *sua*, *meu*, *este*, *a*, *um*, *aquele*, etc. Esses exemplares, a partir de suas regras combinatórias, podem combinar-se com diferentes exemplares da categoria vogal temática do nome (por exemplo *cola*, *pente*, *cabelo*).

Essas DNAs são conhecidas como padrões de DNAs probabilísticos, que configuram exatamente essa possibilidade de variação entre seus elementos dentro de um mesmo conjunto. Van der Bos et al. (2012) investigaram em que medida se dá a diferença de aprendizagem de padrões de DNAs determinísticos e probabilísticos por adultos. Seus resultados mostraram que quando não havia outras pistas visuais ou acústicas (ex. prosódicas), os participantes que foram testados na condição de padrões de DNAs probabilísticos tiveram desempenho bastante inferior quando comparados ao grupo de participantes que foram testados na condição de padrões de DNAs determinísticos. No entanto, nos dois experimentos subsequentes, os autores associaram pistas visuais e acústicas (experimentos 2 e 3, respectivamente) para cada elemento integrante da dependência não-adjacente. Com isso, eles puderam perceber uma melhora significativa do desempenho dos participantes da condição de padrões DNAs probabilísticos comparados aos participantes de padrões de DNAs determinísticas. Os autores concluíram, então, que a aprendizagem de DNAs probabilísticas é tão plausível quanto a aprendizagem de determinísticas, desde que haja pistas adicionais à estrutura.

Em estudos com bebês canadenses (Name et al., 2011; Name & Shi, 2015; Laguardia et al., 2015) e brasileiros de 11 meses (Laguardia et al., 2013; Laguardia et al., 2015), foram usados padrões de DNAs probabilísticos para testar a sensibilidade dos participantes aos estímulos e a capacidade de abstração dessas regras quando em contato com novos exemplares. Foram encontrados resultados significativos nas tarefas de fixação visual, sugerindo que os bebês não só foram sensíveis aos padrões, como também conseguiram abstrair e generalizar as regras a novos exemplares que seguiram os mesmos padrões dos estímulos a que foram expostos.

Um outro exemplo de estudo com padrões de DNAs probabilísticos é o de Henrique (2023). Nesse estudo com adultos, a autora investigou a aprendizagem de DNAs durante um curto tempo de exposição, sem voltar a atenção dos participantes aos estímulos durante a fase de familiarização. Os participantes foram divididos em dois grupos: um experimental e um controle. Os participantes do grupo experimental, durante a fase de familiarização, foram expostos aos estímulos linguísticos do experimento. Esses estímulos eram pseudopalavras e remetiam à DNA presente no português brasileiro encontrada no sintagma determinante (DP), onde para cada dois determinantes distintos havia a possibilidade de duas terminações vocálicas distintas (ex: *koi* e *zis* com terminações em *e* e *u*, *dai* e *nur* com terminações em *o* e *a*)¹². Para o grupo controle, os estímulos foram sons instrumentais de um violão. Os participantes de ambos os grupos foram instruídos a escolher uma imagem abstrata enquanto ouviam os respectivos

¹² Koi kozú, zis vomê; dai dezô, nur fidá

estímulos e a fase teste contou com uma tarefa de escolha forçada, quando ambos os grupos precisavam escolher o padrão probabilístico de DNA congruente com a regra apresentada durante a fase de familiarização para o grupo experimental. Os resultados apontaram para uma aprendizagem dos adultos do grupo experimental desses padrões, além de uma abstração das regras subjacentes quando a acurácia das respostas ficou acima da chance na fase teste para exemplares inéditos do material interveniente.

Uma das questões que essa dissertação levanta no estudo de aprendizagem de DNAs por crianças de 7 anos é que esses padrões de DNAs probabilísticos ainda não foram investigados para essa faixa etária.

Nos dois estudos de base para este presente trabalho (Lammertink et al., 2019; Marimon et al., 2021) que olharam diretamente para a mesma faixa etária, os padrões de DNAs foram determinísticos, com semelhanças e diferenças metodológicas que serão destacadas a seguir.

Lammertink et al. (2019) e Marimon et al. (2021) utilizaram a tarefa de monitoramento de palavra para medir o tempo de resposta (TR) dos participantes ao serem instruídos a pressionar um botão verde (Lammertink et al., 2019) ou branco (Marimon et al., 2021) toda vez que ouvissem a palavra-alvo e um botão vermelho para quando a palavra-alvo não surgisse em uma sequência de três palavras que deveriam ser monitoradas. O trabalho de Lammertink e colegas (2019) utilizou estímulos de uma língua artificial, enquanto o de Marimon e colaboradores (2021) utilizou estímulos de uma língua natural (estímulos em italiano com crianças alemãs monolíngues).

Para o trabalho com língua artificial (Lammertink et al., 2019), as DNAs eram constituídas por pseudopalavras (ex. *tep X lut*) sempre apresentadas em sequências individuais de três palavras, fossem elas na condição gramatical ou agramatical, alvo ou não-alvo. Já para o trabalho de Marimon e colegas (2021), que utilizou língua natural, o padrão escolhido foi uma DNA no VP (*sta Xando* ou *puó Xare*) sempre acompanhado pelo mesmo DP (*Il fratello*).

Em ambos os estudos, para todo elemento *a* individual havia um elemento *b* individual que formava a DNA, não havendo nenhuma variação desses elementos que constituísse uma outra condição gramatical. A partir dos resultados dos TRs dessas tarefas de monitoramento de palavra, os pesquisadores puderam observar uma sensibilidade aos padrões de DNAs utilizados nas tarefas pelas crianças que participaram.

Diante dos estímulos dos estudos citados, onde as DNAs apresentadas aos participantes não apresentaram nenhuma variação entre os elementos, questiona-se se essa aprendizagem implícita observada não passaria apenas de uma aprendizagem condicionada. Dessa forma, não se exclui a aprendizagem estatística envolvida na aprendizagem da dependência apresentada,

mas não se responde se os resultados seriam semelhantes em uma DNA com possíveis variações em seus elementos (a e b) inseridos em classes gramaticais de itens funcionais e governados por regras abstratas. Em outros termos, não se pode falar de aprendizagem de uma regra abstrata, mas sim de uma aprendizagem de um padrão combinatório de elementos em uma configuração aXb .

4.2 O PAPEL DA ATENÇÃO NA INVESTIGAÇÃO DE APRENDIZAGEM DE DNAs

No capítulo 2, seção 2.3, algumas condições foram delineadas em relação à facilitação de aprendizagem de DNAs. Algumas delas foram o número de variabilidade do material interveniente da DNA (Gómez, 2002) e pistas visuais ou acústicas atreladas aos elementos constituintes da DNA probabilística (Van Der Bos et al., 2012).

Em um artigo de revisão, Wilson et al. (2020) além de destacarem esses dois requisitos, também apontaram que a atenção dos participantes ou falantes dispensada aos padrões possui papel facilitador para sua aprendizagem. Um dos trabalhos que os autores destacaram como evidência para essa conclusão foi também o artigo de revisão de Perruchet e Pacton (2006) sobre aprendizagem implícita e aprendizagem estatística. Nele, os autores destacaram a relevância da atenção em contexto de tarefas duplas (em inglês, *dual-task*), quando os participantes precisam recrutar diferentes áreas cerebrais (ex., motoras, cognitivas, de memorização, etc) para dar conta de duas tarefas concomitantemente. Segundo os autores, a aprendizagem do padrão em questão (qualquer que seja sua natureza), acaba sendo comprometida devido ao desvio de atenção aos estímulos, já que outros mecanismos precisam ser ativados para um bom desempenho na tarefa concorrente.

Entretanto, a manipulação da atenção aos estímulos pelos participantes nem sempre precisa ser feita em tarefas duplas. López-Barroso et al. (2016) manipularam a atenção dos participantes adultos em um experimento sobre a aprendizagem de DNAs apenas atribuindo uma palavra-alvo diferente para cada grupo durante a fase de familiarização em uma tarefa de monitoramento de palavra. Nessa tarefa, havia três condições de DNAs distintas. Todos os participantes foram expostos aos três tipos de DNAs (*tagi X sira*, *jupo X runi*, *pine X ladu*). Porém, a cada grupo de participantes foi designada uma palavra-alvo, a qual toda vez que o participante a ouvisse, deveria pressionar o botão esquerdo e, caso a palavra-alvo não estivesse presente na sequência pseudofrasal, ele pressionaria o botão direito. A palavra-alvo era sempre o elemento b (no estudo caracterizado como C) de uma das três condições de DNAs. Durante a fase de familiarização, além das três condições de DNAs, havia também sequências distratoras,

compostas por três elementos X (XXX). Todas as DNAs eram determinísticas. Na fase teste, as sequências distratoras foram substituídas por sequências agramaticais (ex. XXC). Os pesquisadores chamaram essa etapa de fase de teste *online*, porque eles queriam observar se a presença do elemento A anteciparia a presença de C, já que em uma sequência XXC, essa previsão não seria possível. Assim, eles calcularam o tempo de respostas dessas duas condições: a da condição AXC, onde C poderia ser qualquer uma das condições (ou seja, não somente a palavra-alvo do participante) e a condição XXC, onde esse C também poderia ser de qualquer uma das três DNAs. Como resultado, todos os participantes apresentaram tempos menores para as DNAs (AXC) que seguiram as regras da fase de treinamento. Para testar explicitamente se atenção ao elemento final da DNA (A_1XC_1) implicaria em um resultado distinto de aprendizagem em relação ao outro elemento final da DNA que não contivesse a palavra-alvo (ex., A_2XC_2), iniciou-se uma fase teste de julgamento gramatical com sequências gramaticais e agramaticais. As sequências gramaticais contaram com X inéditos e todos os participantes deveriam julgar todas as condições de DNAs, mesmo que algumas delas não contivessem a palavra-alvo a qual eles foram treinados. Nas sequências agramaticais, foram feitos dois tipos de manipulações: um elemento A_1 se relacionando com um C_2 (ex. tagi X runi) e, um elemento C vindo na posição inicial e A vindo na final (ex. sira X tagi). Os participantes obtiveram resultado satisfatório quando as sequências agramaticais inverteram posições (C_1XA_1), porém quando houve apenas violação entre os elementos finais das DNAs (ex. A_1XC_2), os participantes falharam em reconhecer. Esse estudo sugere que, implicitamente, a atenção não parece ser um fator crucial para a aprendizagem. Além disso, a atenção não teve papel fundamental para a categorização das posições dos elementos das DNAs. No entanto, na tarefa explícita em que foi necessário reconhecer os elementos As com seus respectivos Cs, a atenção apresentou efeito marginal.

Outra evidência que corrobora a noção de que a atenção aos estímulos não pareça ter um papel tão crucial é que em estudos com bebês (Name et al., 2011; Name e Shi, 2015; Laguardia et al., 2013; Laguardia et al., 2015), os participantes se mostraram capazes de aprender as DNAs mesmo sob uma condição de escuta passiva, visto que bebês nessa idade não possuem capacidade de manter um foco atencional por muito tempo.

Por outro lado, há uma discussão de que essa habilidade de aprendizagem sob escuta passiva sofre uma queda à medida que ficamos mais velhos (Friederici et al., 2013). Nessa visão, passamos por uma queda gradual de uma aprendizagem mais associativa para uma mais baseada em regras já entre os dois e quatro anos de idade (Paul et al., 2020), que acompanha até os 7 anos de idade quando a maturação do córtex pré-frontal se aproxima mais do cérebro

adulto do que de um bebê (Skeide e Friederici, 2016). Entretanto, no estudo de Friederici et al. (2013) com participantes adultos, no qual foi utilizada uma técnica para minimizar a atividade cerebral do córtex pré-frontal esquerdo responsável pelas funções executivas, que envolvem a atenção e a aprendizagem baseada em regras, os participantes, ainda assim, foram capazes de aprender os padrões de DNAs, a partir de mudanças nos mecanismos subjacentes à aprendizagem. Ou seja, na ausência de recursos de mecanismos encarregados pela aprendizagem baseada em regras, mecanismos que normalmente são recrutados em aprendizagem associativa entram em cena. Isso pode ser uma evidência de que, quando o recurso atencional para as regras é comprometido, outros recursos podem ser empregados para a aprendizagem. Vale lembrar o que vimos no capítulo anterior a respeito de como o foco atencional pode nos levar a ignorar aspectos relevantes mas não tão explícitos de certos estímulos (Thiessen et al., 2016)

Há ainda o estudo de Henrique (2023) que também optou por manipular a atenção dos participantes adultos aos estímulos dos padrões de DNAs. Na fase de familiarização, os dois grupos - experimental e controle - foram instruídos a escolherem entre duas imagens abstratas, a que mais lhe agradava. O grupo controle recebeu como estímulo acústico sons instrumentais de violão, enquanto os participantes do grupo experimental ouviam os estímulos dos padrões de DNAs. A fase teste consistiu em uma tarefa de escolha forçada onde ambos os grupos deveriam decidir, entre um par de palavras exibidos na tela (elementos Xb , onde os X eram inéditos), qual se combinava melhor com o estímulo que ouviam (elemento a). Os participantes foram instruídos a responder intuitivamente, porém o mais rápido possível. O resultado apresentou uma diferença significativa para o grupo experimental, apontando que eles foram capazes de aprender os padrões de DNAs mesmo não tendo sua atenção direcionada voltada para eles durante a fase de familiarização. É importante destacar dois pontos desse estudo em relação à atenção, baseado no que já foi discutido nesta seção.

O primeiro é que a fase de familiarização não constituiu uma tarefa dupla, visto que os participantes não foram instruídos em absolutamente nada em relação aos estímulos acústicos (as DNAs, para o grupo experimental). Além disso, como a atenção dos participantes envolveu processamento visual, é possível que o processamento auditivo não tenha sido comprometido de forma que dois recursos atencionais separados podem ter sido ativados (Jiang e Chun, 2001).

Outra hipótese que Jiang e Chun (2001) levantam sobre recursos atencionais é que se não houver uma carga alta perceptual, o sistema atencional seletivo pode ser mais eficiente e filtrar o que não é relevante. Assim, quando o estudo de Henrique (2023) instrui os participantes a apenas escolherem a imagem que mais lhe agrada, sem precisar se atentar a mais

nenhum outro detalhe, é possível que o sistema atencional do participante possa ter ficado livre, ou de certa forma, não sobrecarregado para perceber o estímulo sonoro dos padrões de DNAs na condição experimental, ainda que não fosse uma ação voluntária ou deliberada.

O segundo ponto é que, de maneira parecida como o estudo de Friederici et al. (2013), porém no âmbito da Psicolinguística, na ausência de uma possibilidade de aprendizagem por regras, ocorreu uma aprendizagem sob escuta passiva, ainda que não seja possível afirmar se foi em caráter associativo.

Nas idades de 7-8 anos de idade, mas não necessariamente em aprendizagem de DNAs e sim, na segmentação de palavras em um contínuo de fala, Saffran, Newport & Barrueco (1997) manipularam a atenção dos participantes ao instruírem que colorissem imagens enquanto um estímulo acústico estava sendo produzido. Os participantes não receberam instrução que deveriam prestar atenção e sequer escutar o que estava sendo tocado. As pesquisadoras observaram uma aprendizagem implícita significativa de palavras em um contínuo de fala, a qual elas atribuíram à probabilidade transicional que resultou em uma aprendizagem estatística.

Em ambos os estudos de base para esta dissertação (Lammertink et al., 2019; Marimon et al. 2021), as crianças participantes foram instruídas a prestarem atenção nos estímulos durante a fase de familiarização, ainda que não fossem informadas sobre qualquer regularidade nas sequências apresentadas. A fase de familiarização (que também configurou a fase teste *online*) consistia em uma tarefa de monitoramento de palavra para ambos os trabalhos. Diferentemente do trabalho de López-Barroso et al. (2016), os participantes foram testados somente nas condições da palavra-alvo. Os pesquisadores interpretaram os resultados de forma que as crianças foram capazes de aprender os padrões de DNAs aos quais foram expostas.

No entanto, por termos uma literatura que aponta uma mudança cognitiva de uma aprendizagem associativa para uma aprendizagem baseada em regras durante essa etapa de desenvolvimento, tendo sua culminância por volta dos 7 anos de idade devido à maturação do córtex pré-frontal, faz-se necessário investigar em que medida a atenção nessa faixa etária está associada ou não à aprendizagem implícita sob uma condição de escuta passiva de padrões de DNAs.

Diante disso, na próxima seção, será discutida a aprendizagem implícita e o papel da tarefa experimental nesse tipo de investigação.

4.3 A INVESTIGAÇÃO DE APRENDIZAGEM IMPLÍCITA EM TAREFAS EXPERIMENTAIS

A aprendizagem implícita é entendida como uma aprendizagem que ocorre sem que o sujeito perceba que está aprendendo algo, mas que de alguma forma, ela causa efeitos posteriores em decisões e ações desse sujeito que pensa estar agindo intuitivamente ou de maneira inconsciente (Cleeremans, 2009). Em outras palavras, o conhecimento implícito pode se apresentar mais tarde como um conhecimento tácito, ou seja, quando um sujeito sabe que sabe de algo, mas não consegue explicar o porquê e como sabe daquilo (Reber, 1989).

Investigar algo que seja implícito mas que ao mesmo tempo já esteja consolidado é desafiador. Isso porque não se pode ter acesso a todos os processos que essa aprendizagem sofreu até se tornar um conhecimento tácito ou mesmo explícito. Assim, em estudos de aquisição de linguagem, com participantes que já tenham uma língua consolidada (crianças mais velhas e adultos, por exemplo), o uso de uma gramática artificial ou pseudonatural pode viabilizar esse tipo de investigação ao selecionar e isolar o objeto de investigação para atestar a existência de uma aprendizagem implícita.

Há um debate na literatura sobre qual a melhor tarefa experimental para testar os participantes sobre o conhecimento implícito ou inconsciente do objeto investigado (Reingold e Merikle, 1988; Cleeremans, 2009; Timmermans e Cleeremans, 2015). Nesse debate, uma das questões levantadas são as tarefas *online* e *offline*¹³ que apresentam medidas indiretas e diretas¹⁴, respectivamente. Porém, Reingold e Merikle (1988) defendem que essas medidas indiretas ou diretas que podem testar conhecimento implícito ou explícito, respectivamente, estão mais associadas ao tipo de instrução e como os estímulos são apresentados, do que à

¹³ Tarefa *online* aqui é entendida como uma tarefa que atesta a aprendizagem durante o processamento dos estímulos. Ex: tarefa de monitoramento de palavra, *Serial Response Time Task*, etc. E tarefa *offline* é entendida como uma tarefa depois que ocorre a aprendizagem, ainda que seja quase que imediatamente após a fase de familiarização. Ex: julgamento gramatical, tarefa de escolha forçada, etc.

¹⁴ Medidas indiretas são entendidas como as respostas dos participantes a partir de uma tarefa experimental, onde eles não são instruídos de antemão sobre a natureza de suas respostas na fase de familiarização. Por sua vez, medidas diretas, são respostas dos participantes que são instruídos ou perguntados explicitamente sobre a natureza dos estímulos. Reingold & Merikle (1988) exemplificam as duas medidas a partir da tarefa de Stroop (1935). Nela, se os participantes foram instruídos a nomearem as palavras na fase teste, quando na fase de familiarização eles também tenham sido instruídos a prestarem atenção nas palavras, então, estamos diante de uma medida direta. No entanto, se na fase de familiarização, os participantes foram instruídos a se atentarem às cores das palavras, para posteriormente precisarem nomear uma palavra, então, estamos diante de uma medida indireta.

designação de uma tarefa em si. Dessa forma, o *design* experimental de uma tarefa, seja ela *online* ou *offline*, não é capaz de determinar o tipo de aprendizagem que está sendo avaliado.

Cleeremans (2009) reforça essa ideia ao defender que a ciência do participante em relação ao caráter de um estímulo, ou até mesmo, a noção de qual é a intenção do pesquisador ou do que vai ser investigado, podem comprometer essa avaliação se, de fato, a tarefa conduz para uma aprendizagem implícita.

Voltemos, então, aos estudos de base deste trabalho para analisarmos como as tarefas experimentais foram conduzidas e como os seus resultados foram interpretados.

Partimos, primeiramente, do estudo de Lammertink et al. (2019) e de suas duas perguntas de pesquisa. A primeira era se era possível medir a sensibilidade de crianças em idade escolar primária às DNAs através dos tempos de resposta (TRs). E a segunda, era medir a sensibilidade de crianças em idade escolar primária às DNAs por meio de tarefa *offline*. Essas perguntas foram suscitadas a partir da ideia de que nessa idade investigada, a técnica de olhar ou escuta preferencial não se aplicaria, já que trata-se de participantes já falantes e com uma L1 consolidada mas que ao mesmo tempo, não possuem uma capacidade metalinguística plenamente desenvolvida (Bialystok, 1986) a ponto de executarem tarefas de julgamento de gramaticalidade que demandam esse tipo de habilidade (ex. escolha forçada, julgamento gramatical, etc.).

Como já descrito em 4.1, a fase de familiarização do experimento de Lammertink e colegas (2019) foi uma tarefa de monitoramento de palavra. Antes que a fase de familiarização se desse início, os participantes foram instruídos a prestarem atenção nas palavras, porque mais tarde perguntas seriam feitas sobre elas. Não foi informado a eles nem se havia regularidade nas sequências de palavras, nem sobre o teor das perguntas que seriam feitas mais tarde. A tarefa foi projetada como um jogo em que as crianças precisavam coletar bananas para o macaco *Appie*. Dessa forma, a cada trinta sequências de três palavras, as crianças recebiam o *feedback* de quantas bananas já haviam coletado.

Houve três blocos de treinamento, seguidos por um bloco de ruptura onde só havia sequências agramaticais, seguida por um bloco de recuperação que configurava uma fase teste (Figura 1).

Figura 1 - Tarefa experimental Monitoramento de Palavras (Lammertink, 2019)

Bloco 1 - treinamento 24 - alvos 24 não alvos 12 distratores	Bloco 2 - treinamento 24 - alvos 24 não alvos 12 distratores	Bloco 3 - treinamento 24 - alvos 24 não alvos 12 distratores	Bloco 4 - ruptura 30 sequências f X b (12 lut e 12 mip)	Bloco 5 - recuperação 24 - alvos 24 não alvos 12 distratores
---	---	---	--	---

Fonte: Elaborada pela autora (2024)

A variável dependente dessa tarefa foram os TRs dos participantes em relação ao terceiro bloco de treinamento mais o de recuperação comparado aos TRs do bloco de ruptura. A previsão era que se as crianças fossem sensíveis à DNA da condição designada, os TRs para os elementos *bs* (na configuração *aXb*) deveriam aumentar no bloco de ruptura, em relação aos dois blocos de treinamento anteriores e diminuir novamente em relação ao bloco de recuperação. De acordo com os autores, os resultados confirmaram essa previsão.

Para responder a segunda pergunta da pesquisa, foi então executada uma tarefa de Escolha Forçada, logo após o primeiro experimento. Os estímulos dessa tarefa foram todos sonoros. Assim, as crianças ouviam duas sequências de três palavras e deveriam responder oralmente qual sequência mais se assemelhava ao padrão apresentado na primeira fase. O resultado ficou no nível da chance e os pesquisadores concluíram que a tarefa de Escolha Forçada não é uma tarefa sensível o suficiente para avaliar a aprendizagem das DNAs por crianças em idade escolar primária, visto que o experimento anterior a esse (o de monitoramento de palavra) apresentou resultados de que essa aprendizagem havia ocorrido.

Passemos para o segundo estudo de base deste trabalho para, então, entrarmos na discussão aqui proposta.

Em Marimon e colaboradores (2021), foram investigados três aspectos sobre a aprendizagem de DNAs. Assim, como objetivos específicos, a pesquisa buscou examinar se (a) a aprendizagem implícita de DNAs pode ser observada por meio de TRs em tarefas *online* ou ativas, (b) se a aprendizagem de DNAs pode ser também medida por uma tarefa de *Stem Completion* (tarefa de completção, similar à tarefa de Escolha Forçada) e, (c) se a aprendizagem por crianças (em quaisquer tarefas) é modulada por idade. Primeiramente, é importante ressaltar que esse estudo investigou além da faixa etária de 7 anos de idade, as idades de 4, 5, 6 e 8 anos de idade. Porém, como essas idades não entram no escopo da presente pesquisa, os dados dos resultados que forem de contrastes entre as idades não serão analisados aqui.

O primeiro experimento, que foi a tarefa de monitoramento de palavra, replicou o *design* experimental de Lammertink et al. (2019), no entanto com estímulos do italiano já descritos em 4.1. As crianças também recebiam *feedback*, que era o número de estrelas coletadas a cada 24 sequências. No início do experimento, anterior à fase de familiarização, as crianças também foram informadas que deveriam ouvir as sequências atentamente e que deveriam responder perguntas no final do experimento.

Ao final do primeiro experimento, as mesmas crianças foram submetidas à tarefa de *Stem Completion*. Elas foram informadas que deveriam escolher um melhor final (elemento *b - ando* ou *are*) para o início da frase que iriam ouvir (elementos *aX*). Também lhes foi dito que se não soubessem a respostas, eles poderiam tentar adivinhá-las.

Como resultado, no primeiro experimento, pode-se observar um aumento significativo dos TRs (uma média de 36 ms mais longo; $p = .00038$) das crianças quando as regras das DNAs não eram apresentadas (durante o bloco de ruptura), levando os pesquisadores a sugerir que houve a aprendizagem das DNAs por todas as crianças. No segundo experimento, no entanto, a acurácia das respostas ficou em 54%, um pouco acima do nível da chance com pontuações individuais variando entre 18,75% a 93,75%.

Os pesquisadores consideraram o segundo experimento mais explícito, portanto, eles sugerem que a tarefa não foi capaz de avaliar uma aprendizagem implícita por parte dos participantes.

4.3.1 Discussão dos Resultados e Análises dos Estudos de Base

Esta subseção propõe-se a discutir os resultados e análises dos experimentos de Tarefa de Escolha Forçada e *Stem Completion* de Lammertink et al. (2019) e Marimon et al. (2021), respectivamente.

Reber (1989), em seu artigo de revisão, que busca definir o conceito de aprendizagem implícita e conhecimento tácito, discorre também sobre como as instruções dadas aos participantes podem influenciar resultados em estudos empíricos que investigam aprendizagem implícita.

Nesse sentido, o autor propõe que quando as instruções são dadas no início da tarefa, há uma grande possibilidade de resultados favoráveis. Se essas instruções forem combinadas com pistas onde os participantes conseguem extrair o que deve ser aprendido, mesmo que os participantes aleguem posteriormente que não perceberam ou encontraram nenhuma regra, os

resultados apontam para uma aprendizagem implícita. Reber então conclui que quando os estímulos fornecem algum tipo de saliência, procurar por regras pode funcionar.

Por outro lado, quando os participantes são expostos a estímulos que não possuem nenhuma pista (por exemplo, Van Der Bos et al., 2012; ver subseção 2.3.1) mas que são informados que existe uma regra, esses participantes podem ser conduzidos a procurar por regras onde não existem ou podem fazer deduções erradas sobre os estímulos. A conclusão de Reber, nesse aspecto, é que não é possível encontrar regras se você não sabe onde procurá-las.

Diante disso, instruções dadas no início do experimento, ou seja, anteriormente à fase de familiarização, podem servir tanto como facilitadoras ou dificultadoras para a aprendizagem.

Entretanto, instruções dadas ao final da fase de familiarização e no início do teste já não trazem efeitos benéficos independentemente do estímulo apresentado, segundo Reber (1989). O autor apresenta duas razões para essa afirmação. A primeira é que, ao dizer ao participante que o estímulo ao que foi exposto apresentava uma regularidade, essa informação pode induzir um processo de metacognição de algo que o participante não tem acesso conscientemente, comprometendo respostas de cunho intuitivo (quando o participante, por exemplo, acredita estar adivinhando ou escolhendo algo que “soa” melhor). O segundo argumento de Reber é que, quando essa atenção é chamada para uma regularidade que até então não existia para o participante, ele se vê com muito menos exemplares para conseguir extrair qualquer regra ou regularidade existente, o que novamente contraria a ideia de que esteja sendo testada uma aprendizagem implícita.

Se analisarmos a condução dos experimentos dos estudos base deste trabalho (Lammertink et al., 2019; Marimon et al. 2021), podemos perceber que há instrução entre as fases de familiarização/tarefa de monitoramento de palavra e a tarefa de Escolha Forçada (*Stem Completion* para Marimon et al., 2021). Em ambos os trabalhos, os participantes são instruídos a responder às tarefas de julgamento de gramaticalidade baseando-se nos estímulos linguísticos da fase anterior.

Dessa forma, os resultados no nível da chance e um pouco acima da chance (Lammertink et al., 2019; Marimon et al. 2021, respectivamente) nessas tarefas de gramaticalidade podem ser analisados como tarefas que investigaram a aprendizagem de DNAs às quais os participantes já poderiam ter algum acesso consciente, e não aprendizagem implícita.

Os pesquisadores avaliam a tarefa de julgamento de gramaticalidade como não sendo a melhor para testar a aprendizagem implícita, quando na verdade, eles desconsideram a noção

de que a resposta intuitiva é um dos fatores que caracterizam esse tipo de aprendizagem à qual o participante ainda não tem um acesso consciente.

Cleeremans (2007) defende que as representações implícitas influenciam no desempenho, porém elas não são suficientemente robustas para o sistema saber que as possui.

O estudo de Henrique (2023) contou com a intuição dos participantes para executarem a tarefa de Escolha Forçada. Após a fase de familiarização (já descrita em 4.1), os participantes foram instruídos a escolherem entre duas palavras (elementos Xb) escritas na tela, a que mais combinava com o estímulo acústico (elemento a). Nenhuma menção foi feita aos estímulos apresentados na fase de familiarização do grupo experimental. Assim, os participantes não tinham nenhum compromisso em lembrar ou acertar algo a que foram expostos anteriormente.

Sobre a noção de intuitividade, Reber (1989) defende que, o participante que acredita estar agindo de forma intuitiva porque algo parece estar certo mesmo se ele não puder verbalizar o motivo, passou por uma experiência de aprendizagem implícita e dessa forma, age sob influência dessa experiência não consciente.

Portanto, quando Henrique (2023) apresenta resultados da aprendizagem de DNAs por adultos com acurácia de 68,01% para o grupo experimental e 45,58% para o grupo controle, estamos diante de uma tarefa de julgamento de gramaticalidade, considerada *offline*, mas que usou medidas indiretas para investigação de aprendizagem implícita. Existe, então, um direcionamento de evidência, de que esse tipo de tarefa possa atestar aprendizagem implícita.

O que permanece uma lacuna ainda nas pesquisas é se essa faixa etária dos 7 anos, que foi investigada pelos estudos de base deste trabalho (Lammertink et al., 2019; Marimon et al. 2021) também se mostra capaz de apresentar resultados significativos de aprendizagem implícita de padrões de DNAs em tarefa de julgamento de gramaticalidade, quando direcionada para a intuição das respostas.

Para isso, aplicaremos o experimento de Henrique (2023) com pequenas adaptações para o público alvo, que serão descritas e discutidas no próximo capítulo, a fim de explorar todas essas questões levantadas neste capítulo.

3.5 CONCLUSÃO

Nas seções anteriores, foram apresentados o referencial teórico que sustenta as discussões propostas nesta dissertação, assim como os trabalhos base, que contemplam a mesma faixa etária, e seus desdobramentos.

Apresentamos os padrões de DNAs determinísticos e probabilísticos e como eles podem guiar a investigação da aprendizagem de DNAs para diferentes análises, tendo as DNAs determinísticas como ferramenta para avaliação de sensibilidade dos participantes aos padrões, e as probabilísticas como fontes de investigação para abstração e generalização de regras. Vimos ainda que para a faixa etária de 7 anos de idade, há dados de sensibilidade aos padrões de DNAs determinísticos, mas que ainda existe a lacuna para os padrões probabilísticos.

Olhamos como a atenção tem sido manipulada em estudos sobre a aprendizagem de DNAs em diferentes faixas etárias e quais os resultados desses estudos. Também vimos como os recursos atencionais podem ser poupados ou subutilizados quando uma tarefa concomitante ou usa mecanismos distintos (motores, cognitivos, etc.) ou envolve processamentos separados (ex., auditório e visual), ou não causa uma carga perceptual pesada em uma das tarefas. Na faixa etária deste trabalho, apresentamos evidência de resultado positivo significativo em tarefa de segmentação de palavras quando a atenção não estava voltada aos estímulos, o que pode ser um indicativo que essa idade pode aprender sob condição de escuta passiva. Em relação à aprendizagem sob escuta passiva, apresentamos estudos que defendem que existe uma mudança gradual no tipo de aprendizagem à medida que nos desenvolvemos cognitivamente, mas que ao mesmo tempo, na ausência de possibilidade de aprendizagem baseada em regras, outros mecanismos podem tomar a frente, viabilizando a aprendizagem dos padrões apresentados. Levantamos a questão, então, se a escuta passiva e a ausência de atenção aos estímulos por crianças de 7 anos pode ser uma condição plausível para a aprendizagem implícita de padrões de DNAs.

Em relação à aprendizagem implícita, defendemos a ideia de que a condução da tarefa experimental de Escolha Forçada diz mais sobre o tipo de investigação do que a denominação da tarefa *per se*. Para isso, trouxemos os estudos de Reber (1989) que discutem como e quando as instruções dadas aos participantes podem apresentar resultados favoráveis ou desfavoráveis à pesquisa. Também, ressaltamos o quão imprescindível deve ser a noção de resposta intuitiva dada pelos participantes ao se investigar a aprendizagem implícita. E por fim, discutimos as análises dos resultados apresentados pelos estudos de base deste trabalho (Lammertink et al., 2019; Marimon et al. 2021), quando os autores indicam que a tarefa de Escolha Forçada não parece ser uma tarefa sensível para investigação de aprendizagem implícita de DNAs em crianças de 7 anos de idade.

5 ATIVIDADE EXPERIMENTAL

Neste capítulo, teremos uma breve apresentação sobre a Tarefa de Escolha Forçada, suas características e as razões de sua escolha para este projeto. Em seguida, será apresentada a proposta de experimento que foi desenvolvido, com seus objetivos, hipóteses, variáveis dependente e independentes, condições experimentais e previsões, além do perfil dos participantes e os tipos de estímulos que foram utilizados. Ainda, apresentaremos em detalhes todas as etapas que a atividade experimental contemplou. Por fim, serão apresentados os resultados, incluindo suas análises e articulação com os estudos de base deste trabalho.

5.1 A TAREFA DE ESCOLHA FORÇADA

No último capítulo, na seção 4.3, argumentamos sobre a possibilidade da tarefa de Escolha Forçada ser uma técnica experimental eficiente na investigação de aprendizagem implícita de DNAs por crianças com pelo menos uma língua já consolidada.

Nesta seção, veremos como normalmente essa técnica se apresenta e quais as variáveis dependentes que são extraídas dela.

A tarefa de Escolha Forçada, ou em inglês, *Two-forced Alternative Task* (2AFC), em alguns estudos de Psicolinguística, se apresenta como uma tarefa de julgamento de gramaticalidade. Tarefa de julgamento de gramaticalidade é aquela em que os participantes julgam, de forma consciente ou não, se um estímulo é compatível ou não com os estímulos linguísticos apresentados durante a fase de familiarização e/ou com seu conhecimento de uma língua, L1 ou L2. No entanto, a tarefa de Escolha Forçada nada mais é que, uma tarefa onde os participantes devem, necessariamente escolher entre uma das duas opções, qualquer que seja a natureza dos estímulos que estão sendo avaliados.

Especificamente sobre a tarefa de Escolha Forçada, o participante deve escolher sempre, entre duas opções, aquela que atende ao requisito da tarefa. Por exemplo, após os participantes serem expostos a estímulos na fase de familiarização, durante a fase de teste, eles podem ser requisitados a responder sim ou não para sequências que eles lembram ter visto ou ouvido na fase anterior. Outra possibilidade é quando as duas alternativas são itens ou partes de palavras que devem se concatenar a um outro item apresentado no início de cada *trial* do teste.

Essa tarefa possibilita a utilização tanto de estímulos sonoros quanto de visuais. Isso vale para ambas as fases: a de familiarização e a de teste. Existe, ainda, a possibilidade de

mesclar os dois tipos de estímulos durante as fases. Um exemplo é quando, na fase teste, o item a ser concatenado aparece na forma sonora e as duas opções nas formas visuais, ou na configuração contrária.

A tarefa de Escolha Forçada é considerada uma tarefa *offline*, ou seja, ela não capta o momento em que a aprendizagem ou o processamento dos estímulos ocorre, mas o resultado dessa aprendizagem ou processamento, seja ela implícita(o) ou explícita(o). Além disso, essa tarefa pode utilizar medidas diretas ou indiretas - como já explicadas na seção 4.3 do capítulo anterior - que, em outras palavras, avaliam respectivamente a aprendizagem (resultante de um processamento) sobre o que os participantes pensavam estar aprendendo, ou algo subjacente aos estímulos apresentados (Reingold & Merikle, 1988).

As variáveis dependentes que costumam ser extraídas desse tipo de tarefa são os tempos de resposta (TR) e a acurácia das respostas. Essas duas variáveis combinadas direcionam para uma análise de resultados de que se houve ou não a aprendizagem dos padrões apresentados nos estímulos. Enquanto um curto tempo de resposta pode indicar uma tomada de decisão rápida e confiante, um índice de acurácia acima do nível da chance pode indicar que as respostas não foram dadas de forma arbitrária.

Os estudos de base para este trabalho (Lammertink et al., 2019; Marimon et al. 2021) utilizaram essa técnica experimental com crianças de idades entre 4 e 8 anos. Como já foi discutido na seção 4.3 do capítulo anterior, existem algumas questões deixadas em aberto relativas à aplicação da tarefa experimental nessa faixa etária na investigação de aprendizagem implícita de DNAs.

Assim, como Henrique (2023) obteve resultados significativos com adultos ao aplicar a tarefa de Escolha Forçada, recorrendo à intuitividade dos participantes em suas respostas, decidimos aplicar seu experimento, fazendo algumas adaptações necessárias para atender as demandas motivacionais de crianças de 7 anos de idade. Entendemos que as crianças deveriam ser envolvidas em uma narrativa em que eles tivessem uma participação ativa. A simples justificativa de que eles participariam de uma atividade experimental sobre aprendizagem de língua não parecia empolgante o suficiente para eles se sentirem tão engajados. Eles foram informados que participariam de uma pesquisa acadêmica, mas que também ajudariam um personagem criado especialmente para essa pesquisa. Todas as etapas e adaptações serão descritas nas seções e subseções a seguir.

5.2 EXPERIMENTO

Como o experimento foi realizado com crianças, o contato inicial foi feito com o responsável do participante. Após a apresentação das assinaturas do Termo de Consentimento Livre e Esclarecido (TCLE) do responsável pelo participante, o responsável respondeu, juntamente com o participante, um questionário¹⁵ sobre fluência, hábito de leitura e informações pessoais (idade, ano escolar, diagnóstico de neuro divergências, contato com L2, etc.) do participante que seriam relevantes para a análise de resultados e que foi entregue à experimentadora antes que o experimento se iniciasse.

A experimentadora, então, apresentava a narrativa do experimento. Era apresentada uma folha A4¹⁶ com o desenho de um robô e duas opções de nomes para os participantes escolherem: *Rob* e *Bot*. A experimentadora explicava que o robô era muito indeciso e que iria precisar da ajuda da criança para tomar algumas decisões, a começar pelo seu nome. Assim, a criança escrevia seu nome abaixo do nome da sua escolha e a partir daí, o experimentador tratava o robô pelo nome escolhido pela criança.

Antes que o experimento desse início, a experimentadora pedia aos participantes que lessem algumas palavras. Essas palavras faziam parte de um teste de leitura que consistia de, em uma tela, duas palavras que seguiam o *layout* (Figura 2) da fase teste da atividade experimental. Foram apresentados três *slides* com duas palavras cada. Do total de 6 palavras, três foram palavras do Português Brasileiro (PB) e três foram pseudopalavras.

Todas as palavras atenderam aos critérios dos estímulos que seriam apresentados na fase de familiarização e teste: palavras dissílabas e oxítonas. A configuração apresentada nos slides foi : (i) duas palavras¹⁷ do PB, (ii) uma palavra do PB e uma pseudopalavra e (iii) duas pseudopalavras.

Os participantes leram as palavras em voz alta para a experimentadora. A intenção desse teste era que a experimentadora certificasse que o participante fosse capaz de ler as palavras com uma certa fluência e em um curto espaço de tempo, além de dar a ênfase correta na sílaba tônica. É importante lembrar que crianças de 7 anos ainda estão passando por um processo de alfabetização e assim, somente o questionário respondido pelos responsáveis pode não ser uma fonte inteiramente precisa para um bom desempenho na tarefa a ser realizada. Somente um

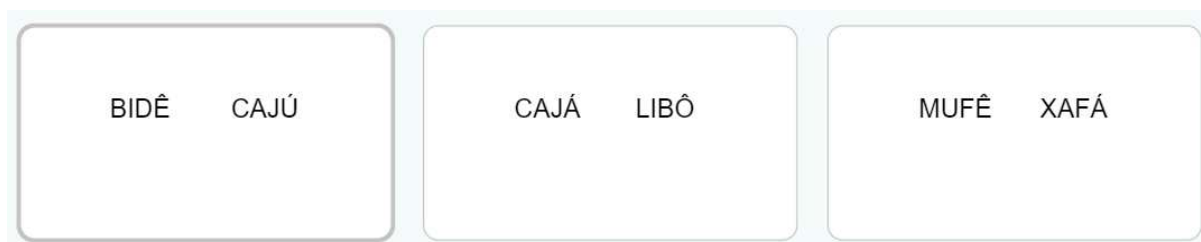
¹⁵ Apêndice A

¹⁶ Apêndice B

¹⁷ O acento gráfico na palavra “caju” é meramente para criar uma homogeneidade na forma de apresentação das palavras e pseudopalavras do teste de leitura, assim como as pseudopalavras que aparecem na tela para os participantes durante a fase de teste.

participante não passou no teste de leitura por ainda não estar alfabetizado. Ele participou do experimento mesmo assim, visto que a exclusão de sua participação poderia causar uma frustração à criança, porém os dados não foram utilizados.

Figura 2 - *Slides* do teste de leitura



Fonte: Elaborado pela autora (2024)

Encerrado o teste de leitura, a primeira fase do experimento foi descrita e explicada ao participante, para que quaisquer dúvidas pudessem ser esclarecidas antes do início do processo. Como mencionado na seção 4.1 deste capítulo, aplicamos a atividade experimental de Henrique (2023). A fim de trazer ludicidade e, consequentemente, engajar as crianças na atividade, foi dito aos participantes que eles precisariam ajudar o robô (Figura 3) em mais uma decisão. A decisão contemplou a fase de familiarização que será explicada a seguir, porém como narrativa, foi dito que o robô tinha acabado de se mudar para uma casa cheia de paredes brancas e que queria muitos quadros para enfeitar sua galeria. Como o robô era muito indeciso, precisava da ajuda das crianças para escolher o quadro mais bonito. Elas também foram informadas que usariam um fone de ouvido, por conta de um áudio durante a tarefa. No entanto, nada foi informado aos participantes sobre o conteúdo do áudio e se eles deveriam prestar atenção nele.

Figura 3 - Tela apresentada aos participantes antes do início da atividade



Fonte: Elaborada pela autora (2024)

O experimento contou com duas fases e em ambas as fases, a apresentação dos estímulos foi programada de forma aleatória por meio da plataforma Farm PCIBex¹⁸. Sobre o procedimento das fases, tivemos:

1. Fase de Familiarização

Durante a fase de familiarização, os participantes foram divididos aleatoriamente em dois grupos: o grupo experimental e o grupo controle. A todos eles, foi informado que o robô precisava escolher alguns quadros para sua nova casa. Assim, todos os participantes foram instruídos a escolher uma imagem, entre as duas que apareceriam na tela (Figura 4), a que mais lhe agradasse. Para a imagem da esquerda, eles deveriam pressionar a tecla C do teclado e para a imagem da direita, eles deveriam pressionar a tecla M. Como adaptação para os participantes, as teclas C e M tiveram um adesivo azul colado nas teclas para melhor visualização.

O grupo experimental, enquanto executava a tarefa de escolha das imagens, escutou 128 sequências de padrões linguísticos constituídos por DNAs, descritos em 5.2.2.2. O grupo controle, por sua vez, como estímulo acústico, ouviu uma sequência de sons instrumentais de um violão, também de autoria de Henrique (2023). A tarefa de familiarização teve uma duração variável, já que o tempo dependia do tempo de escolha das imagens, podendo chegar até 3

¹⁸ O site Farm PCIBex (<https://farm.pcibex.net/>) foi desenvolvido por Zehr & Schwarz (2018) na Universidade da Pensilvânia, e fornece ferramentas para criação e execução de experimentos na área da Psicolinguística.

minutos. Os dados e análise desse tempo de duração para ambos os grupos serão apresentados na subseção 5.2.3.3.

Figura 4 - Exemplo das imagens durante a fase de familiarização



Fonte: Henrique, 2023

2. Fase teste

Após a fase de familiarização, todos os participantes (grupos experimental e controle) seguiram para a fase teste. Antes do início dessa fase, a experimentadora contou para cada participante a outra parte da narrativa. Foi dito a eles que estava sendo criada uma nova língua para o robô e que esse robô gostaria da ajuda do participante para escolher as palavras que melhor se combinavam nessa nova língua.

Após o participante sinalizar para a experimentadora que entendeu a tarefa, a experimentadora dava início à fase teste.

Na fase de teste, os participantes tiveram que executar uma tarefa de Escolha Forçada. Nessa fase, todos os participantes foram expostos aos mesmos estímulos, sem distinção de grupos.

Após apresentado um estímulo acústico (Figura 5) que poderia um dos quatro determinantes (KOI/ZIS/DAI/NUR) que foram apresentados ao grupo experimental durante a fase de familiarização, o participante deveria decidir entre duas opções de pseudonomes inéditos¹⁹ que estavam escritos na tela e apresentavam o mesmo padrão silábico e fonotático da fase de familiarização. Eles deveriam pressionar as mesmas teclas sinalizadas com adesivos azuis na fase de familiarização. Se achassem que a palavra da esquerda combinasse melhor com o estímulo auditório, deveriam pressionar a tecla à esquerda, caso a palavra escolhida estivesse à direita, pressionariam a tecla à direita. Essa decisão ou escolha foi em caráter

¹⁹ Ver anexo A.

intuitivo, visto que os participantes não foram informados sobre a possibilidade de haver uma resposta incorreta. Foram apresentados 16 pares de palavras para esta tarefa, e ela teve duração de aproximadamente 2 minutos.

Figura 5 - Exemplo do *layout* da tela durante a fase de teste



Fonte: Henrique, 2023

Após a realização dessa etapa, o experimento chegava ao fim. Cada criança recebeu um certificado simbólico de participação por ter ajudado o robô.

Diante do exposto, trabalharemos a partir do objetivo, definições das variáveis independentes e dependente, hipótese e previsões.

5.2.1 Objetivo

O objetivo deste trabalho é investigar se crianças aos sete anos de idade são sensíveis às DNAs contidas em conjuntos de uma classe morfológica de uma língua pseudonatural, por um tempo curto de exposição.

5.2.2 Método

Foram duas variáveis independentes: (i) a linguística foi o padrão de DNA entre determinantes e vogais temáticas (A(Det) X B(-vogal temática)) na pseudolíngua. Essa variável apresentou dois níveis: congruência ou incongruência entre Det e terminação de N de acordo com o que foi apresentado na fase de familiarização para o grupo experimental; e (ii) como

fator grupal, *familiarização*, também com dois níveis: estímulo linguístico (grupo experimental) e música (grupo controle).

A variável dependente é a acurácia da resposta. No estudo original de Henrique (2023), a pesquisadora também calculou o tempo de resposta dos participantes nas escolhas dos itens na fase teste. No entanto, ela não encontrou nenhuma diferença estatisticamente significativa entre as escolhas de itens congruentes ou incongruentes na fase teste, em ambos os grupos. Diante desses resultados, não nos pareceu pertinente usar o TR como variável dependente. Um outro fator que contribuiu para essa decisão, foi o fato das crianças aos 7 anos de idade ainda estarem passando pelo processo de alfabetização, o que poderia impactar no tempo individual necessário para a leitura em cada *trial*, comprometendo uma análise mais criteriosa.

Dessa forma, propomos a hipótese de que crianças de 7 anos de idade são capazes de aprender implicitamente padrões de DNAs probabilísticos.

Assim, nossas previsões são:

(1) Se as crianças do grupo experimental forem capazes de aprender implicitamente padrões de DNAs probabilísticos, a acurácia de suas respostas deverá ser acima do nível da chance.

(2) O grupo controle, que não foi exposto aos estímulos linguísticos durante a familiarização, deverá ter taxa de acurácia de suas respostas baixa ou no nível da chance.

Em suma, buscamos atingir os objetivos específicos da seguinte maneira:

(i) verificar, a partir de uma tarefa experimental, se crianças aos sete anos de idade são sensíveis às DNAs contidas em conjuntos de uma classe morfológica de uma língua pseudonatural, mesmo que por um tempo curto de exposição:

- utilizando uma parte de pseudolíngua que simula a DNA no PB (A(Det) X B(-vogal temática)), onde dois pares de determinante podem ser combinados com duas vogais temáticas (ex., koi/zis X-e/u);
- formando um grupo experimental e outro controle. Apresentando os estímulos linguísticos durante a fase de familiarização somente para os participantes do grupo experimental;
- e expondo os participantes aos estímulos linguísticos do grupo experimental durante a fase de familiarização por, no máximo, 3 minutos.

(ii) investigar em que medida a atenção direcionada aos padrões linguísticos apresentados (DNAs) é requisito para sua aprendizagem implícita nessa faixa etária:

- direcionando a atenção de todos os participantes (grupo experimental e controle) durante a fase de familiarização, ao instruí-los a escolher uma imagem entre duas;

- não mencionando ou instruindo sobre os estímulos sonoros (estímulos linguísticos e sons de violão);
 - levando todos os participantes a acreditarem, durante a fase teste, que eles estão fazendo uma escolha pessoal e intuitiva sobre as combinações das palavras (DNAs).
- (iii) avaliar a eficácia a Tarefa de Escolha Forçada como um método experimental na investigação de aprendizagem de DNAs por crianças de 7 anos de idade, analisando possíveis necessidades de modificações e adaptações:
- engajando os participantes, em uma narrativa lúdica, a fazerem escolhas a partir de suas preferências pessoais, sugerindo que as respostas de ambos os grupos sejam de cunho intuitivo.
 - verificando se a taxa de acurácia para os participantes do grupo experimental se mostra acima do nível da chance e apresenta uma diferença estatisticamente significativa em comparação à taxa do grupo controle.

5.2.2.1 Participantes

Participaram do estudo 35 crianças. Desse número, 3 participantes não foram incluídos nas análises dos resultados. Um participante por não passar no teste de leitura e outros dois por precisarem de assistência durante o experimento, podendo, assim, comprometer os resultados obtidos. Portanto, ao final, tivemos um número de 32 crianças, que foram divididas em igual número em dois grupos: 16 no grupo experimental (meninos N= 9) e 16 no grupo controle (meninos N= 7). Os participantes tinham idade entre 6 anos, 8 meses e 4 dias e 8 anos, 0 meses e 8 dias ($M = 7a3m17d$). Não houve nenhum critério de exclusão anterior ao experimento, além do fator idade. Participaram crianças de escolas particulares e públicas, incluindo algumas cujos responsáveis alegaram alguma neurodivergência. Foi estipulado que a exclusão só seria feita após o experimento, caso o participante não conseguisse seguir os passos das tarefas sem interferências, visto que as crianças mostravam interesse em participar. As crianças que tinham idade adjacentes aos 7 anos de idade, foram aceitas por estarem cursando o mesmo ano e estarem na mesma sala de aula de outros participantes deste estudo.

O recrutamento dos participantes foi feito através de redes sociais e aplicativos de conversas. As crianças ou responsáveis não receberam nenhum tipo de gratificação ou colaboração financeira pela participação na atividade, apenas um certificado de participação, em forma de agradecimento.

Os responsáveis foram informados sobre a aprovação deste projeto pelo Comitê de Ética da UFJF²⁰, e foram solicitados a assinarem, caso estivessem de acordo, o Termo de Consentimento Livre e Esclarecido (TCLE). Além disso, os responsáveis foram solicitados a responderem o questionário sobre as crianças.

5.2.2.2 Materiais utilizados

Toda a atividade experimental foi executada em um *laptop* HP Pavillion RT3290 e um fone de ouvido infantil *Headphone Oex Hp 305*. Como apoio, foi utilizada uma mesa para o *laptop*, além de uma cadeira para o participante se sentar.

Todos os estímulos foram retirados de Henrique (2023). Não foi feita nenhuma alteração em relação às pseudopalavras, visto que era de interesse deste estudo a comparação entre os dados a serem obtidos e os resultados obtidos pela autora do experimento. A autora defendeu o uso de uma pseudolíngua, argumentando que seu uso é uma boa maneira de minimizar interferências semânticas e fonéticas nos estímulos, quando a investigação é feita em participantes com uma L1 já consolidada.

Os padrões de DNAs utilizados neste trabalho foram probabilísticos. Assim, os determinantes KOI e ZIS poderiam ser combinados com nomes dissílabos e oxítonos terminados em [e] ou [u]. Da mesma maneira, os outros dois determinantes, DAI e NUR, poderiam ser combinados com os nomes dissílabos e oxítonos terminados em [a] e [o] (Quadro 1). A escolha por palavras oxítonas também foi uma tentativa de distanciar um pouco do padrão dos substantivos do PB, visto que, em sua maioria, são palavras paroxítonas. As consoantes iniciais e mediais dessas pseudopalavras foram balanceadas, assim como as vogais presentes em suas primeiras sílabas²¹. Segunda a autora, os determinantes deste experimento foram criados cuidadosamente para evitar possíveis semelhanças com padrões fonéticos de determinantes no Português. Foram apresentadas 128 sequências dessas palavras para o grupo experimental durante a fase de familiarização.

Na seção 2.3 e nas suas respectivas subseções, descrevemos algumas pistas adicionais que podem facilitar a aprendizagem de DNAs. Dessa forma, nossos estímulos continham algumas dessas pistas. A primeira delas é que eles foram apresentados entre pausas e não em um contínuo de fala, sendo possível o participante perceber onde uma DNA começava e terminava. A outra, é o fato de serem palavras oxítonas. Embora essa tonicidade fuja do padrão

²⁰ Parecer no.6.967.869

²¹ Ver anexo B

dominante do PB, a acentuação tônica na última sílaba apresenta uma saliência no elemento *b* da DNA.

Em relação à variabilidade do X, para cada par de determinante (ex. koi/zis , dai/nur), foram apresentados durante a fase de familiarização 32 elementos X, 16 concatenados com uma das possíveis combinações (ex. 32 X, 16 com terminações -e, 16 com terminação -u, para os determinantes koi e zis). Sobre o princípio de Gestalt, proposto por Newport et al. (2004), como um facilitador na aprendizagem de DNA, os determinantes KOI e DAI (cada um pertencendo a um par distinto), atendem à similaridade fonética com as terminações vocálicas e/u e a/o, respectivamente. Por outro lado, os determinantes ZIS e NUR não cumprem esse princípio, já que eles possuem uma terminação consonantal.

Quadro 1 - Exemplos de padrões dos estímulos acústicos experimentais

DETERMINANTE	KOI / ZIS	DAI / NUR
TERMINAÇÃO FINAL DO NOME	[e] [u]	[a] [o]
Exemplos de estímulos	Koi [3e'tu] Koi [la'ne] Zis [mu'ze] Zis [xo'vu]	Dai [pi'fo] Dai [te'ba] Nur [nu'bo] Nur [zu'va]

Fonte: Henrique, 2023

5.2.2.3 O ambiente

Durante a realização da atividade, além do participante, esteve presente a experimentadora. Se o responsável pelo participante optasse por estar presente durante a atividade, ele poderia fazê-lo. A execução da atividade, por não envolver muitos aparatos, pôde ser feita por vezes na residência do participante, além do laboratório do NEALP, localizado no Centro de Humanidades e Pesquisas, na Universidade Federal de Juiz de Fora.

A condição do ambiente para a execução da tarefa foi um ambiente mais silencioso possível, com baixa interferência de outros estímulos.

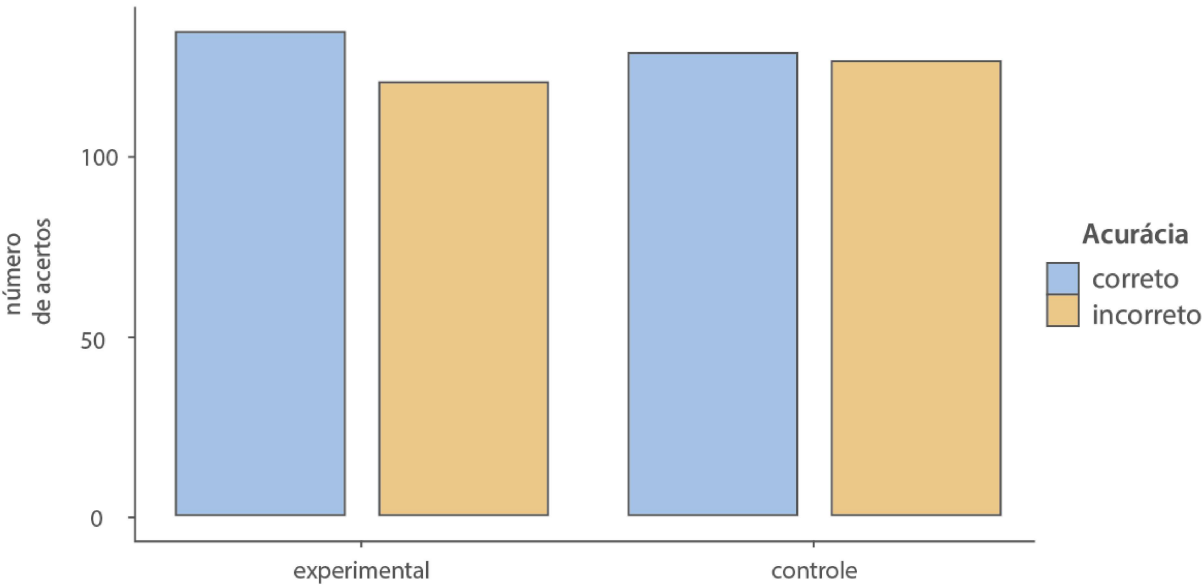
5.2.3 Análise e discussão

Nesta subseção, apresentaremos os resultados da aplicação da atividade experimental, assim como as respectivas análises estatísticas. A variável dependente analisada foi a acurácia das respostas de todos os participantes durante a fase teste da atividade experimental. A acurácia foi determinada a partir da congruência entre a escolha de uma DNA da fase teste e uma DNA apresentada na fase de familiarização. Em seguida, compartilharemos algumas análises exploratórias a fim de ampliar a discussão sobre o método experimental aplicado, os dados obtidos através dele e condições que possam ter influenciado ou não no resultado final. Ao final, ainda como análise exploratória, teremos a relação de tempo de familiarização de ambos os grupos e algumas interpretações e direcionamentos a partir desses dados. E, por fim, discutiremos nossos resultados com os resultados obtidos nos estudos de base deste trabalho.

5.2.3.1 Resultados

Os participantes do grupo experimental, que foram apresentados aos estímulos linguísticos durante a fase de familiarização, apresentaram um resultado um pouco acima da chance de acertos nas escolhas das respostas congruentes durante a fase teste. Com um total de 256 itens (16 pares de itens para 16 participantes), o número de acertos foi de 135 (52,7%), enquanto o de erro foi de 121 (47,2%). O grupo controle, cujos participantes não foram expostos a nenhum estímulo linguístico na fase de familiarização, apresentou um número próximo de acertos e erros durante a fase teste para a escolha de sequências linguísticas. Também com dezesseis pares de itens para cada participante (N=16) escolher durante a fase teste, totalizando 256 respostas, o grupo controle fez 129 (50,3%) escolhas corretas, comparado a 127 (49,6%) escolhas incorretas, apresentando um resultado no nível da chance.

Gráfico 1 - Acertos e erros cometidos pelos participantes dos grupos experimental e controle durante a fase teste



Fonte: Elaborado pela autora, 2024

No entanto, apesar dos participantes do grupo experimental terem apresentado um número maior de acertos de respostas na fase teste em relação aos participantes do grupo controle, quando comparados estatisticamente, através de um teste χ^2 , essa diferença não foi significativa ($\chi^2 = 0.282$; $p = 0.596$).

A partir da tabela (Tabela 1) abaixo, é possível ver a distribuição de acertos por participantes em cada condição. Dessa forma, podemos perceber como as médias de acertos se apresentam de maneira bem semelhante entre ambos os grupos.

Tabela 1 - Média de acertos e erros pelos participantes de ambos os grupos durante a fase teste.

Estatística Descritiva				
	Condição	N	Média	Desvio-padrão
Acertos	Controle	16	8.06	1.53
	Experimental	16	8.44	2.13
Erros	Controle	16	7.94	1.53
	Experimental	16	7.56	2.13

Fonte: Elaborado pela autora, 2024

Portanto, diante destes resultados e suas respectivas análises, não podemos afirmar que houve aprendizagem implícita dos padrões de DNAs pelos participantes do grupo experimental. Embora as taxas de acertos mais altas tenham advindo dos participantes do grupo experimental, as médias de acertos são praticamente iguais entre ambos os grupos, principalmente quando levamos em conta o desvio padrão. Na próxima subseção, ampliaremos a discussão desses resultados e fatores que devem ser levados em consideração.

5.2.3.2 Análises Exploratórias

Apesar de não termos obtido resultados estatisticamente significativos, algumas análises exploratórias foram feitas, a fim de verificar alguns fatores que possam ter influenciado os resultados.

Os fatores que decidimos explorar no grupo experimental (*within subjects*), em relação à taxa de acertos, foram: (1) idade dos participantes; (2) tempo de familiarização; (3) sexo dos participantes e; (4) ano escolar em que o participantes estavam cursando durante a coleta de dados. A seguir, apresentaremos as justificativas para essas análises e seus respectivos resultados. Contudo, é importante ressaltar que o número de participantes investigados (N=16) pode ser uma amostra muito pequena para qualquer tipo de conclusão. Por serem análises exploratórias, advindas de questionamentos dos nossos resultados, achamos conveniente investigar alguma tendência de interação entre os quesitos analisados, e no caso de um resultado positivo, essa interação poderia guiar estudos futuros nesta linha de pesquisa. Para isso, foi feito o teste χ^2 para cada um desses fatores.

O primeiro fator analisado, idade dos participantes, foi considerado pelo fato desta faixa etária de 7 anos de idade passar por processos de maturação cerebral, cognitivos, metacognitivos, de aprendizagem e de educação formal (alfabetização), como visto na seção 3.2. Assim, era importante observar se havia alguma relação entre uma idade mais avançada com uma maior taxa de acertos pelos participantes durante a fase teste. Dividimos os participantes do grupo experimental em dois grupos, os com idade abaixo de 7 anos e 6 meses (N= 8) e os com idade acima deste corte. A percentagem de acertos para as crianças abaixo de 7 anos e meio foi de 49,21%, enquanto para as crianças de idade acima de 7 anos e 6 meses foi de 56,25% (ver tabela 2 para números de acertos e erros de cada grupo). Em um teste χ^2 , não foram encontradas diferenças significativas entre os grupos de menor ou maior idade que 7 anos e seis meses em relação à acurácia das respostas ($\chi^2= 1.27, p= 0.260$). Dessa forma, com

o número de participantes disponíveis para esta análise, não podemos afirmar que o fator idade pareça ter influência relevante dentro deste estudo para uma maior acurácia das respostas.

Tabela 2 - Tabela de contingência relacionando a acurácia das respostas pelos participantes do grupo experimental em relação às suas respectivas idades.

Tabela de contingência			
	acurácia		
idade	incorreto	correto	Total
> 7.5	56	72	128
< 7.5	65	63	128
Total	121	135	256

Fonte: Elaborado pela autora, 2025

O segundo fator a ser analisado foi a relação entre a taxa de acurácia e o tempo de familiarização, à qual os participantes do grupo experimental foram submetidos. Essa análise se fez relevante porque o tempo máximo de exposição aos estímulos linguísticos durante a fase de familiarização pelos participantes do grupo experimental era de três minutos, podendo ser um pouco mais curto a depender do tempo de escolha das imagens nessa fase. Ou seja, à medida que os participantes escolhiam uma imagem do par, uma nova sequência do padrão de DNA se dava início. Dessa forma, tornou-se necessário explorar se os participantes do grupo experimental que tiveram um tempo mais curto de exposição durante essa fase, tiveram menor desempenho na fase teste quando comparados aos participantes que utilizaram o tempo total ou quase total de familiarização.

Para essa observação, também foi utilizado um teste χ^2 entre o número de respostas corretas na fase teste e o tempo de familiarização divididos em dois grupos (< 2.669 e > 2.669), calculado a partir da média do tempo de familiarização de todos os participantes do grupo experimental(tempo mínimo de 2.03 minutos e tempo máximo de 3 minutos), com 7 participantes pertencendo ao grupo de menor tempo e 9 no de maior tempo. Abaixo, podemos observar os resultados obtidos dessa análise.

Tabela 3 - Taxa de acurácia das respostas pelos participantes do grupo experimental em relação ao tempo de familiarização (máximo de 3 minutos).

Tabela de contingência			
	acurácia		
tempo de familiarização	incorreto	correto	Total
> 2.6	46	50	96
< 2.6	75	85	160
Total	121	135	256

Fonte: Elaborado pela autora, 2025

A partir da Tabela 3, podemos perceber que a média de tempo de familiarização para os acertos, não se difere tanto para os erros. Ou seja, a proporção de acertos para os participantes que tiveram mais tempo de exposição aos estímulos linguísticos durante a fase de familiarização (52,08%) não foi maior do que os participantes que tiveram uma exposição um pouco mais curta durante esse mesmo período (53,12%). A análise estatística do teste χ^2 ($\chi^2 = 0.261$, $p = 0.872$), também demonstra que essa diferença não é significativa. Portanto, não é possível afirmar que a diferença encontrada entre tempos de familiarização maiores ou menores dos participantes do grupo experimental e seus respectivos números de respostas corretas são diretamente proporcionais.

O terceiro fator a ser explorado foi o sexo dos participantes (meninos N= 9 e meninas N= 7). Embora essa questão não tenha sido preditora sobre uma diferença de desempenho em nenhuma literatura referenciada nesta dissertação, por ser um fator de diferenciação entre os participantes, decidimos por explorar a possibilidade de um dos sexos ter taxas de acertos superiores ao outro. Contudo, como era esperado, não foi encontrada nenhuma diferença significativa ($\chi^2 = 3.81$, $p = 0.149$) no desempenho entre meninos e meninas no grupo experimental.

Tabela 4 - Média de desempenho dos participantes do grupo experimental em relação ao sexo.

Estatística Descritiva				
	Sexo	N	Média	Desvio-padrão
taxa de acertos	menino	9	8.78	2.54
	menina	7	8.00	1.53

Fonte: Elaborado pela autora, 2024

O quarto, e último fator a ser explorado, foi o ano escolar (1º ano N = 9 e 2º ano N= 7) em que o participante estivesse cursando durante a coleta de dados. Embora o resultado da relação da idade dos participantes com o número de acertos das respostas já poderia ser um indicativo, a escolha por essa análise exploratória se deu ao fato de que esses dois anos escolares acompanham o processo pedagógico de alfabetização das crianças. Dessa forma, se o nível de proficiência na leitura dos itens durante a fase teste fosse um ponto crucial para um bom desempenho, o ano escolar poderia ser um fator de grande influência nesse quesito. Assim, também foi feito um teste χ^2 , relacionando a taxa de acertos das respostas de cada participante na fase teste e o ano escolar em que cada um deles estava cursando.

Como pode ser observado abaixo (Tabela 5), os alunos do segundo ano do fundamental alcançaram uma média mais alta do que os alunos do primeiro ano. Porém, ao observarmos o desvio-padrão, essa diferença não se mostra relevante. Estatisticamente, as médias das taxas de acertos também não apresentam diferença significativa ($\chi^2 = 2.96, p = 0.227$).

Tabela 5 - Média de acertos dos participantes do grupo experimental em relação ao ano escolar.

Estatística Descritiva				
	ano escolar	N	Média	Desvio-padrão
taxa de acertos	2º ano	7	9.29	1.98
	1º ano	9	7.78	2.11

Fonte: Elaborado pela autora, 2024

5.2.3.3 Tempo de familiarização para os grupos experimental e controle

Em nosso *design* experimental, os grupos experimental e controle receberam estímulos auditivos distintos durante a fase de familiarização. Enquanto os participantes do grupo experimental ouviram estímulos linguísticos dos padrões probabilísticos das DNAs, os participantes do grupo controle, ouviram sons instrumentais de um violão. Ambos os grupos receberam o mesmo tipo de instrução durante essa fase: eles deveriam escolher, entre duas imagens, aquela que mais gostassem. A única menção sobre o estímulo auditivo era que eles precisavam usar um fone de ouvido, já que iriam escutar um áudio durante a tarefa. Para nenhum dos grupos, o teor do áudio foi enfatizado. Durante a aplicação do experimento, no entanto, a experimentadora observou que, alguns participantes do grupo controle executavam a tarefa (especialmente a tarefa de escolha de imagem) com muito mais rapidez do que os participantes do grupo experimental.

Como análise exploratória, decidimos investigar se essa diferença também era estatisticamente significativa e o que ela poderia nos indicar.

Abaixo, podemos observar as médias, tempos mínimos e máximos da fase de familiarização para ambas as condições (Tabela 6), além do teste estatístico χ^2 ($\chi^2 = 4.80$; $p = 0.028$) (Tabela 7).

Com isso, é possível notar uma diferença significativa entre os tempos de familiarização para as duas condições grupais. Em um dos nossos objetivos específicos, questionamos em que medida a atenção direcionada aos padrões linguísticos apresentados (DNAs) seria requisito para a aprendizagem implícita na faixa etária de 7 anos de idade. Por não termos encontrado diferenças estatisticamente significativas na taxa de acurácia da fase teste entre os grupos experimental e controle, não é possível responder, a partir dessa diferença no tempo de familiarização, sobre o papel na aprendizagem. No entanto, dado o fato de que os dois grupos receberam o mesmo tipo de instrução durante a tarefa, além de terem o foco de sua atenção desviada dos estímulos auditivos (linguísticos ou não), essa diferença estatística pode ser um indicativo interessante.

Como essa análise não faz parte do escopo do presente trabalho, não cabe a discussão exaustiva desses resultados. Porém, esses dados podem indicar que, mesmo quando a atenção não é direcionada para os estímulos linguísticos, parece que alguma parte da atenção é utilizada para essa percepção, ainda que involuntária. No próximo capítulo, discutiremos, ainda que brevemente, o que essa análise pode representar e como essa questão poderia ser investigada futuramente.

Tabela 6 - Médias, tempos mínimos e máximos durante a fase de familiarização para ambos os grupos

Estatística Descritiva						
	Condição	N	Média	Desvio-padrão	Mínimo	Máximo
Tempo de familiarização	Controle	16	2.13	0.778	0.986	3.00
	Experimental	16	2.67	0.339	2.030	3.00

Fonte: Elaborado pela autora, 2025

Tabela 7 - Teste χ^2 para o tempo de familiarização em ambos os grupos

Testes χ^2			
	Valor	gl	p
χ^2	4.80	1	0.028
N	32		

Fonte: Elaborado pela autora, 2025

5.2.3.4 Discussão

No presente estudo, buscamos investigar se crianças de 7 anos de idade são capazes de aprender implicitamente dependências não adjacentes probabilísticas. Como objetivos específicos, observamos, a partir de uma tarefa experimental de Escolha Forçada, se crianças nessa faixa etária seriam sensíveis aos dois pares de DNAs probabilísticas, às quais elas foram expostas por um curto período de tempo. Além disso, tentamos verificar em que medida a atenção direcionada aos estímulos linguísticos seria crucial para a aprendizagem dos padrões em questão.

Nossos resultados não apontaram diferenças estatisticamente significativas na acurácia das respostas da tarefa de Escolha Forçada entres os participantes do grupo experimental e controle. Portanto, não é possível sugerir que houve aprendizagem implícita dos padrões de DNAs pelos participantes do grupo experimental.

Algumas das questões suscitadas durante o desenvolvimento desta dissertação partiram de dois trabalhos, que adotamos como estudo de base, sobre aprendizagem de DNAs com

participantes da mesma faixa etária de 7 anos. Dessa forma, a comparação entre os resultados encontrados nesses trabalhos com os nossos permite aprofundar em algumas discussões.

Como vimos na subseção 4.3.1, a execução de tarefa de Escolha Forçada nesses dois trabalhos pode ter comprometido a investigação da aprendizagem implícita dos participantes, visto que a eles foram dadas instruções explícitas sobre o que deveria ser levado em conta durante a escolha do item durante a fase teste. Em nossa atividade experimental, baseada no trabalho de Henrique (2023), tomamos o cuidado para que as respostas dos participantes fossem de cunho intuitivo, para que, caso o resultado fosse significativo, pudéssemos argumentar que havia ocorrido uma aprendizagem implícita. Diante de tudo isso, quando observamos os resultados do grupo e individuais da tarefa de Escolha Forçada de Lammertink et al. (2019) e Marimon et al. (2021), comparando-os com os nossos (Tabela 8), levantamos algumas questões.

Tabela 8 - Comparação entre os resultados obtidos no presente trabalho e os resultados dos estudos de base

	Acurácia do grupo	Acurácia individual
Lammertink et al. (2019)	51,6 %	entre 25 e 75%
Marimon et al. (2021)	54,94%	entre 18.75 e 93.75%
Fernandes (2025)	52,7%	entre 25 e 68,75%

Fonte: Elaborado pela autora, 2025

A primeira delas é que mesmo com a atenção desviada dos estímulos linguísticos, os nossos participantes tiveram desempenho muito parecido com os participantes dos dois outros trabalhos que participaram ativamente de uma tarefa *online* anterior à fase teste, onde sua atenção era voltada a pelo menos um padrão de DNA. Isso pode indicar que a atenção direcionada aos estímulos linguísticos não foi uma facilitadora para um melhor desempenho dos participantes dos estudos de base, em relação aos nossos participantes, já que as taxas de acurácia encontradas nos três estudos são bastante próximas.

A segunda questão é em relação ao grau de dificuldade da tarefa. Os participantes dos estudos de base contaram apenas com dois pares de DNAs determinísticas²², enquanto os nossos participantes foram expostos a dois pares de DNAs probabilísticas²³. Apresentando

²² Lammertink et al. (2019) - *tep X lut; sot X mip* e Marimon et al. (2021) *sta Xand/puó Xare*.

²³ Henrique (2023) - *koi* ou *zis X* [e] ou [u]; *dai* ou *nur X* [a] ou [o]

desempenho parecido, o aumento do número de DNAs a ser aprendido não foi inversamente proporcional ao resultado do desempenho, o que pode indicar um grau de dificuldade similar para a aprendizagem de DNAs determinísticas e probabilísticas.

Por último, devido à aplicação da tarefa experimental de Henrique (2023), é natural compararmos nossos resultados com os da autora. Em seu trabalho com adultos, a pesquisadora encontrou como resultado, 68.01% de acurácia das respostas dos participantes. Um número bastante expressivo que sugere a aprendizagem implícita das DNAs pelos adultos. A questão que surge dessa comparação, adicionado aos resultados dos trabalhos de base é se, crianças de 7 anos de idade não são capazes de aprender implicitamente padrões de DNAs, sejam eles determinísticos ou probabilísticos, ou se, como tarefa experimental, a literatura ainda carece de uma técnica experimental mais sensível e adaptada para essa faixa etária que forneça a taxa de acurácia das respostas como variável dependente. Essa indagação vem somada ao fato de que os estudos de base encontraram resultados significativos de aprendizagem de DNA com crianças de 7 anos em uma tarefa *online* (monitoramento de palavras) em que a variável dependente foi o tempo de resposta. Dessa forma, nossa hipótese de trabalho prevaleceria, ou seja, crianças de 7 anos de idade são capazes de aprender implicitamente padrões de DNAs probabilísticos.

Uma outra interpretação alternativa, visto que manipulamos a atenção dos participantes dos estímulos linguísticos durante a fase de familiarização, é de que essa manipulação possa ter comprometido a aprendizagem e, consequentemente, os resultados encontrados. Como não tivemos um terceiro grupo que tivesse sua atenção voltada a esses estímulos linguísticos, não temos propriedade para afirmar se a atenção desviada, de fato, foi a responsável pela baixa taxa de acurácia encontrada. No entanto, essa é uma questão que merece ser investigada com mais cuidado futuramente.

Timmermans e Cleeremans (2015) ressaltam a importância de se avaliar cada decisão metodológica tomada em estudos de aprendizagem implícita. Segundo os autores, pequenas diferenças na aplicação do experimento podem levar a uma gama de conclusões distintas. Além disso, eles reforçam a importância dos pesquisadores se apoiarem em um rico arcabouço teórico para que todas as decisões sobre a metodologia sejam respaldadas por estudos anteriores. Dessa forma, a partir de um referencial teórico aqui apresentado, com diversos estudos sobre aprendizagem de DNAs descritos a partir de suas tarefas experimentais, além dos estudos de base (Lammertink et al., 2019; Marimon et al., 2021) e o estudo original (Henrique, 2023), cujo experimento foi aqui adaptado e aplicado, podemos afirmar que as decisões metodológicas adotadas foram todas baseadas na investigação de aprendizagem estatística e

implícita, que contemplou a faixa etária de 7 anos de idade. Em outras palavras, qualquer alteração na tarefa experimental proposta neste trabalho poderia configurar uma investigação de outros tipos de aprendizagem (ex: explícita ou metalinguística). Por outro lado, concordamos com Timmermans e Cleeremans (2015) quando assumimos que nenhum método pode ser considerado puro a ponto de não sofrer nenhuma interferência. Portanto, há sempre espaço para aprimorarmos técnicas e *designs* experimentais, a fim de buscarmos resultados mais esclarecedores.

No próximo capítulo, apresentaremos alguns direcionamentos para pesquisas futuras, a partir das questões aqui levantadas.

5.3 CONCLUSÃO

Neste capítulo, descrevemos brevemente as características da tarefa de Escolha Forçada e como ela costuma ser utilizada em estudos empíricos. Em seguida, apresentamos nossa atividade experimental baseada no trabalho de Henrique (2023), detalhando as adaptações para o público alvo e as fases contempladas pelo experimento.

Retomamos nossos objetivos geral e específicos, relatando as ações necessárias para atingir cada um deles. Apresentamos as características de nossos participantes e nossos estímulos linguísticos utilizados.

Na subseção de análise e discussão, apresentamos os resultados de acertos e erros dos grupos experimental e controle na fase teste de nossa atividade experimental. Vimos que, apesar dos participantes do grupo experimental terem atingido um número maior de acertos em comparação com o grupo controle, essa diferença não foi estatisticamente significativa. Portanto, não se pode afirmar que houve aprendizagem implícita dos padrões de DNAs pelas crianças do grupo experimental.

Embora não tenhamos encontrado um resultado significativo, trouxemos algumas análises exploratórias para certificarmos de que outros fatores pudessem ter comprometido os resultados obtidos. Assim, em uma configuração de análises intra-sujeitos (*within subjects*), relacionamos as taxas individuais de acurácia dos participantes do grupo experimental aos seguintes fatores: (1) idade dos participantes, (2) tempo de familiarização dos participantes, (3) sexo e, (4) ano escolar que o participante estava cursando durante a coleta de dados. Em nenhuma dessas análises foi encontrada diferença significativa, validando os resultados obtidos através da nossa atividade experimental.

No entanto, em uma outra análise exploratória, dessa vez inter-sujeitos (*between subjects*), foi encontrada uma diferença significativa em relação ao tempo de familiarização entre os participantes do grupo experimental e do grupo controle. Essa diferença chamou a atenção pelo fato dos dois grupos terem recebido instruções idênticas mas apresentarem comportamentos distintos diante de somente uma variável, o estímulo acústico. Sendo assim, participantes que receberam estímulos linguísticos durante a fase de familiarização (experimental) utilizaram mais tempo na tarefa do que aqueles que não receberam o mesmo tipo de estímulo (controle). Essa análise abre uma questionamento para o papel da atenção não direcionada do participante quando um estímulo linguístico está em segundo plano.

Na subseção de discussão, comparamos nossos resultados com os estudos de base deste trabalho (Lammertink et al., 2019; Marimon et al., 2021). A partir dessa comparação, duas questões foram levantadas: (i) que o fato dos participantes do estudo de base terem o foco de sua atenção voltada para os estímulos linguísticos não deu nenhuma vantagem a eles no desempenho na tarefa de Escolha Forçada em relação aos nossos participantes e; (ii) que nossos participantes, que foram testados com DNAs probabilísticas, apresentaram resultados semelhantes aos participantes dos estudos de base, que trabalharam com DNAs determinísticas, acendendo uma discussão sobre a aprendizagem de DNAs em relação ao grau de dificuldade nessa faixa etária. Além dos estudos de base, também olhamos para os resultados de Henrique (2023), a fim de refletir sobre a grande diferença de resultado com o nosso trabalho, juntamente com as informações do estudo de base. Isso nos levou a questionar se a faixa etária dos 7 anos de idade (e idades adjacentes) conta com uma atividade experimental sensível o suficiente para avaliar aprendizagem estatística, especificamente, aprendizagem de DNAs, tendo como variável dependente a taxa de acurácia das respostas.

No próximo capítulo, serão apresentadas considerações finais sobre as questões suscitadas neste trabalho, assim como, as contribuições ao estudo de aquisição de linguagem e os direcionamentos de pesquisas futuras a partir desse apanhado.

6. CONSIDERAÇÕES FINAIS

Este trabalho propôs-se a investigar a aprendizagem implícita de dependências não adjacentes por crianças de 7 anos. Como objetivos específicos, buscamos (i) observar, a partir de uma tarefa experimental, se crianças aos sete anos de idade são sensíveis às DNAs contidas em conjuntos de uma classe morfológica de uma língua pseudonatural, mesmo que por um tempo curto de exposição; (ii) investigar em que medida a atenção direcionada aos padrões linguísticos apresentados (DNAs) é requisito para sua aprendizagem implícita nessa faixa etária; (iii) e avaliar a eficácia da Tarefa de Escolha Forçada como um método experimental na investigação de aprendizagem de DNAs por crianças de 7 anos de idade, analisando possíveis necessidades de modificações e adaptações. Trinta e duas crianças, com média de idade de 7 anos, 3 meses e 17 dias, participaram da atividade experimental, divididas em dois grupos, um experimental e um controle.

Como resultado, não foi registrada uma diferença estatisticamente significativa da acurácia das respostas na fase teste das crianças que formaram o grupo experimental, das que formaram o controle. Dessa forma, a partir de nosso primeiro objetivo específico, os resultados sugerem que crianças aos sete anos de idade não se mostraram sensíveis às DNAs contidas em classes morfológicas, pelo menos, quando aplicada a tarefa de Escolha Forçada.

Quanto ao segundo objetivo específico, examinamos em que medida a atenção direcionada aos estímulos linguísticos seria um requisito para a aprendizagem implícita aos 7 anos de idade e, para isso, desviamos a atenção dos participantes durante a exposição dos estímulos acústicos (linguísticos para o grupo experimental, não-linguístico para o controle) na fase de familiarização. Como nosso resultado não indicou uma aprendizagem pelos participantes do grupo experimental, ainda não é possível assumir um papel desse tipo de atenção para esse tipo de aprendizagem. Como pesquisa futura, a criação de um terceiro grupo de participantes, que tivessem sua atenção voltada para os estímulos linguísticos de forma indireta poderia ser um caminho para uma melhor investigação sobre o papel da atenção na aprendizagem de DNAs por crianças dessa faixa etária. Outra questão levantada neste trabalho em relação à atenção a partir de uma análise exploratória dos resultados, foi a observação de que os participantes do grupo experimental passaram mais tempo ouvindo os estímulos linguísticos durante a fase de familiarização, comparados com os participantes do grupo controle, que ouviram estímulos não-linguísticos. A diferença desse tempo foi estatisticamente significativa e chama a atenção pelo fato de ambos os grupos terem recebido o mesmo tipo de instrução e o tempo máximo de 3 minutos para executar a tarefa durante essa fase. Portanto,

apesar de não cumprirmos com esse segundo objetivo, essa informação pode ser um direcionamento para estudos futuros que investiguem o papel da atenção não direcionada dos participantes quando se tratar de estímulos linguísticos e sua aprendizagem.

Sobre o terceiro objetivo específico, considerando que dedicamo-nos a sanar alguns problemas e questões levantadas sobre os estudos de base na execução da tarefa de Escolha Forçada, podemos entender que esse objetivo foi cumprido. Como contribuição, este trabalho sugere que a técnica de Escolha Forçada para atestar aprendizagem de DNAs em crianças na faixa etária de 7 anos de idade pode não ser sensível o suficiente para resultados comprobatórios e satisfatórios. Visto que é uma atividade experimental amplamente utilizada em estudos empíricos, com bons resultados em adultos (ex., López-Barroso, 2016; Henrique, 2023), ao passo que com bebês, a técnica de fixação visual parece cumprir o mesmo papel (ex., Name et al., 2011; Name e Shi, 2015; Laguardia et al., 2013; Laguardia et al., 2015), é possível que estejamos diante de uma lacuna no que diz respeito à uma técnica experimental voltada para uma faixa etária que esteja passando por um grande desenvolvimento cerebral, cognitivo e metacognitivo, ou seja, entre os 5 e 10 anos (McNealy et al., 2011; Thiessen et al., 2016; Lammertink et al., 2019; Marimon et al. 2021). Assim, parece pertinente que estudos empíricos futuros avaliem novas possibilidades de técnicas experimentais que se voltem para essa fase de desenvolvimento tão importante e impactante para a aprendizagem, e consequentemente, para a aquisição de linguagem.

Este trabalho também buscou contribuir com uma revisão de literatura sistemática sobre a aprendizagem estatística e habilidade computacional, estudos sobre a aprendizagem de DNAs e estudos realizados com a faixa etária aqui investigada em diferentes aspectos, contemplando habilidades distintas (funções executivas, controle inibitório, aprendizagem estatísticas, metacognição, etc.). Dessa forma, espera-se que esta dissertação sirva para consultas futuras, para que mais estudos sobre aprendizagem implícita de DNAs por crianças sejam desenvolvidos a fim de preencher todas essas lacunas aqui descritas.

REFERÊNCIAS

- ASLIN, Richard N.; SAFFRAN, Jenny R.; NEWPORT, Elissa L. Statistical learning in linguistic and nonlinguistic domains. **The emergence of language**, p. 359-380, 1999.
- BARRERA, Sylvia Domingos; MALUF, Maria Regina. Consciência metalingüística e alfabetização: um estudo com crianças da primeira série do ensino fundamental. **Psicologia: reflexão e crítica**, v. 16, p. 491-502, 2003.
- BIALYSTOK, Ellen. Factors in the growth of linguistic awareness. **Child development**, p. 498-510, 1986.
- CHOMSKY, Noam. **The Architecture of Language**. Oxford: Oxford University Press, p. 11-52, 2000.
- CLEEREMANS, Axel. Consciousness: the radical plasticity thesis. **Progress in brain research**, v. 168, p. 19-33, 2007.
- CLEEREMANS, A. Implicit learning and implicit memory. **Encyclopedia of consciousness**, p. 369-381, 2009.
- DEOCAMPO, Joanne A.; KING, Tricia Z.; CONWAY, Christopher M. Concurrent learning of adjacent and nonadjacent dependencies in visuo-spatial and visuo-verbal sequences. **Frontiers in psychology**, v. 10, p. 1107, 2019.
- DE VRIES, Meinou H. et al. Processing multiple non-adjacent dependencies: Evidence from sequence learning. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 367, n. 1598, p. 2065-2076, 2012.
- ENDRESS, Ansgar D.; NESPOR, Marina; MEHLER, Jacques. Perceptual and memory constraints on language acquisition. **Trends in cognitive sciences**, v. 13, n. 8, p. 348-353, 2009.
- EDWARDS, Harold T.; KIRKPATRICK, Anita G. Metalinguistic awareness in children: A developmental progression. **Journal of psycholinguistic research**, v. 28, p. 313-329, 1999.
- FRIEDERICI, Angela D. et al. Language learning without control: the role of the PFC. **Journal of cognitive neuroscience**, v. 25, n. 5, p. 814-821, 2013.
- FROST, Ram et al. Domain generality versus modality specificity: The paradox of statistical learning. **Trends in cognitive sciences**, v. 19, n. 3, p. 117-125, 2015.
- FROST, Rebecca LA; MONAGHAN, Padraic. Simultaneous segmentation and generalisation of non-adjacent dependencies from continuous speech. **Cognition**, v. 147, p. 70-74, 2016.
- FROST, Rebecca LA et al. Non-adjacent dependency learning in infancy, and its link to language development. **Cognitive Psychology**, v. 120, p. 101291, 2020.
- GÓMEZ, Rebecca L. Variability and detection of invariant structure. **Psychological Science**, v. 13, n. 5, p. 431-436, 2002.

GÓMEZ, Rebecca; MAYE, Jessica. The developmental trajectory of nonadjacent dependency learning. **Infancy**, v. 7, n. 2, p. 183-206, 2005.

GRAMA, Ileana; WIJNEN, Frank. Learning and generalizing non-adjacent dependencies in 18-month-olds: A mechanism for language acquisition?. **Plos one**, v. 13, n. 10, p. e0204481, 2018.

HENRIQUE, Késsia da Silva. **Processamento e aprendizagem de dependências não adjacentes por adultos falantes do português brasileiro**. 2023. Tese (Doutorado em Psicolinguística) - Faculdade de Letras, Universidade Federal de Juiz de Fora, 2023.

HUTTENLOCHER, Peter R. Morphometric study of human cerebral cortex development. **Neuropsychologia**, v. 28, n. 6, p. 517-527, 1990.

JIANG, Yuhong; CHUN, Marvin M. Selective attention modulates implicit learning. **The Quarterly Journal of Experimental Psychology: Section A**, v. 54, n. 4, p. 1105-1124, 2001.

LAGUARDIA, M.C.C.T., NAME, C.; SHI, R. Sensitivity to non-adjacent dependencies and grammatical category relations in 11-month-old infants. In: BUCLD 38 The 38th Annual Boston University Conference on Language Development, 2013, Boston. Proceedings of the 38th Boston University Conference on Language Development. Boston, 2013.

LAGUARDIA, M. C. C. T. ; SANTOS, E. ; SHI, R. ; NAME, C. Eleven-month-old infants use prosodic boundaries to learn non-adjacent grammatical dependencies. In: 40th Boston University Conference on Language Development, 2015, Boston. Proceedings of 40th BUCLD, 2015.

LAGUARDIA, Milene Cristine de Castro Teixeira. **Etapas iniciais de aquisição lexical - Habilidades estatísticas e simbólicas no tratamento de dependências não adjacentes**. Tese (Doutorado em Psicolinguística) - Faculdade de Letras, Universidade Federal de Juiz de Fora, 2016.

LAMMERTINK, Imme et al. Auditory statistical learning in children: Novel insights from an online measure. **Applied Psycholinguistics**, v. 40, n. 2, p. 279-302, 2019.

LÓPEZ-BARROSO, Diana et al. Attentional effects on rule extraction and consolidation from speech. **Cognition**, v. 152, p. 61-69, 2016.

MARCUS, Gary F. et al. Rule learning by seven-month-old infants. **Science**, v. 283, n. 5398, p. 77-80, 1999.

MARIMON, Mireia et al. Children's Learning of Non-adjacent Dependencies Using a Web-Based Computer Game Setting. **Frontiers in Psychology**, v. 12, p. 734877, 2021.

MCNEALY, Kristin; MAZZIOTTA, John C.; DAPRETTO, Mirella. Age and experience shape developmental changes in the neural basis of language-related learning. **Developmental science**, v. 14, n. 6, p. 1261-1282, 2011.

NAME, C.; SHI, R.; KOULAGUINA, E.; Non-adjacent dependency learning and grammatical categorization in 11- month-old infants. 12th International Congress for the Study of Child Language. Montreal, 2011.

NAME, C.; SHI, R. Preverbal infants track and represent non-adjacent dependencies at an abstract level. In: 40th Boston University Conference on Language Development, 2015, Boston. Proceedings of 40th BUCLD, 2015.

NEWPORT, Elissa L.; ASLIN, Richard N. Learning at a distance I. Statistical learning of non-adjacent dependencies. **Cognitive psychology**, v. 48, n. 2, p. 127-162, 2004.

ONNIS, Luca et al. Variability is the spice of learning, and a crucial ingredient for detecting and generalizing in nonadjacent dependencies. In: **Proceedings of the annual meeting of the cognitive science society**. 2004.

ONNIS, Luca et al. Phonology impacts segmentation in online speech processing. **Journal of Memory and Language**, v. 53, n. 2, p. 225-237, 2005.

PAUL, Mariella et al. Gradual development of non-adjacent dependency learning during early childhood. **Developmental cognitive neuroscience**, v. 50, p. 100975, 2021.

PEÑA, Marcela et al. Signal-driven computations in speech processing. **Science**, v. 298, n. 5593, p. 604-607, 2002.

PERRUCHET, Pierre; PACTON, Sebastien. Implicit learning and statistical learning: One phenomenon, two approaches. **Trends in cognitive sciences**, v. 10, n. 5, p. 233-238, 2006.

RABI, Rahel; MINDA, John Paul. Rule-based category learning in children: The role of age and executive functioning. **PloS one**, v. 9, n. 1, p. e85316, 2014.

REBER, Arthur S. Implicit learning and tacit knowledge. **Journal of experimental psychology: General**, v. 118, n. 3, p. 219, 1989.

REBER, Arthur S.; ALLEN, Rhianon. Individual differences in implicit learning: Implications for the evolution of consciousness. In: **Individual differences in conscious experience**. John Benjamins Publishing Company, 2008. p. 227-248.

REINGOLD, Eyal M.; MERIKLE, Philip M. Using direct and indirect measures to study perception without awareness. **Perception & psychophysics**, v. 44, n. 6, p. 563-575, 1988.

ROARK, Casey L.; HOLT, Lori L. Auditory information-integration category learning in young children and adults. **Journal of experimental child psychology**, v. 188, p. 104673, 2019.

SAFFRAN, Jenny R.; ASLIN, Richard N.; NEWPORT, Elissa L. Statistical learning by 8-month-old infants. **Science**, v. 274, n. 5294, p. 1926-1928, 1996.

SAFFRAN, Jenny R. et al. Incidental language learning: Listening (and learning) out of the corner of your ear. **Psychological science**, v. 8, n. 2, p. 101-105, 1997.

SCHAADT, Gesa et al. Seven-year-olds' recollection of non-adjacent dependencies after overnight retention. **bioRxiv**, p. 670257, 2019.

SIEGELMAN, Noam; FROST, Ram. Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. **Journal of memory and language**, v. 81, p. 105-120, 2015.

SKEIDE, Michael A.; FRIEDERICI, Angela D. The ontogeny of the cortical language network. **Nature reviews neuroscience**, v. 17, n. 5, p. 323-332, 2016.

TEIXEIRA, Milene Cristine de Castro. **A contribuição das informações distribucionais de DNA nas etapas iniciais da aquisição lexical**. Dissertação (Mestrado em Psicolinguística) - Faculdade de Letras, Universidade Federal de Juiz de Fora, 2012.

THIESSEN, Erik D.; GIRARD, Sandrine; ERICKSON, Lucy C. Statistical learning and the critical period: how a continuous learning mechanism can give rise to discontinuous learning. **Wiley Interdisciplinary Reviews: Cognitive Science**, v. 7, n. 4, p. 276-288, 2016.

TIMMERMANS, Bert; CLEEREMANS, Axel. How can we measure awareness? An overview of current methods. **Behavioural methods in consciousness research**, v. 21, p. 21-46, 2015.

UDDEN, Julia et al. Implicit acquisition of grammars with crossed and nested non-adjacent dependencies: Investigating the push-down stack model. **Cognitive Science**, v. 36, n. 6, p. 1078-1101, 2012.

VAN DEN BOS, Esther; CHRISTIANSEN, Morten H.; MISYAK, Jennifer B. Statistical learning of probabilistic nonadjacent dependencies by multiple-cue integration. **Journal of Memory and Language**, v. 67, n. 4, p. 507-520, 2012.

VAN DER KANT, Anne et al. Linguistic and non-linguistic non-adjacent dependency learning in early development. **Developmental cognitive neuroscience**, v. 45, p. 100819, 2020.

VAN HEUGTEN, Marieke; SHI, Rushen. Infants' sensitivity to non-adjacent dependencies across phonological phrase boundaries. **The Journal of the Acoustical Society of America**, v. 128, n. 5, p. EL223-EL228, 2010.

VINTER, Annie; PERRUCHET, Pierre. Implicit learning in children is not related to age: Evidence from drawing behavior. **Child Development**, v. 71, n. 5, p. 1223-1240, 2000.

WANG, Felix Hao; MINTZ, Toben H. Learning nonadjacent dependencies embedded in sentences of an artificial language: When learning breaks down. **Journal of Experimental Psychology: Learning, Memory, and Cognition**, v. 44, n. 4, p. 604, 2018.

WILSON, Benjamin et al. Non-adjacent dependency learning in humans and other animals. **Topics in cognitive science**, v. 12, n. 3, p. 843-858, 2020.

APÊNDICES

Apêndice A - Questionário sobre as crianças em relação às suas habilidades e hábitos de leitura



Prezada / Prezado Responsável,

O presente questionário tem o objetivo de conhecer melhor a criança, no que se refere a características e hábitos relacionados à pesquisa da qual ele/ela participará. Você tem o direito de não responder qualquer questão, caso você não se sinta à vontade. É importante frisar que esse questionário não possui caráter avaliativo e nenhuma resposta implicará em nenhum tipo de julgamento.

Ainda, reforçamos nosso compromisso de manter todas as informações em sigilo.

Agradecemos seu interesse e sua disponibilidade e a de seu/sua filho/filha em participar da pesquisa.

Nome da criança participante:

Data de nascimento da criança participante:

Nome do/da responsável pela criança:

1. Qual ano escolar seu/sua filho/filha está cursando neste momento?

2. Seu/sua filho/filha possui diagnóstico de dislexia?

☐ sim

☐ não

3. Seu/sua filho/filha é considerado diagnosticamente neurodivergente?

☐ sim Qual o diagnóstico? _____

☐ não

4. Seu/sua filho/a tem o hábito de ler fora do ambiente escolar como forma de entretenimento?

☐ nunca

☐ às vezes

☐ frequentemente

5. Você considera que seu/sua filho/filha tem facilidade para ler textos impressos em diferentes fontes?

a) palito ou fôrma

☐ sim

☐ não

☐ não sei

b) cursiva

☐ sim

☐ não

☐ não sei

c) de imprensa

☐ sim

☐ não

☐ não sei

6. Seu/sua filho/filha costuma fazer as pausas de acordo com pontuação (vírgulas, ponto final, etc) durante a leitura?

☐ nunca

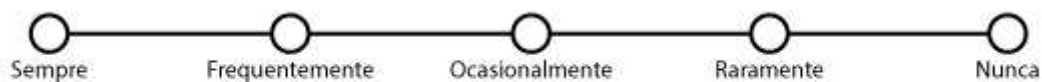
- ☐ às vezes
- ☐ frequentemente

7. Seu/sua filho/filha faz uso de aparelho auditivo?

- ☐ sim
- ☐ não

8. Sobre o contato com uma segunda língua, selecione na escala o cenário mais dentro da realidade de seu/sua filha/filho:

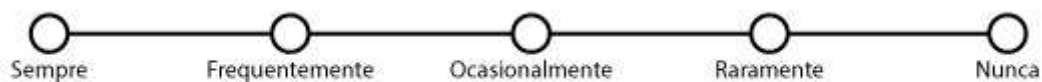
a. Tem aulas extracurriculares de uma segunda língua por duas ou mais horas por semana.



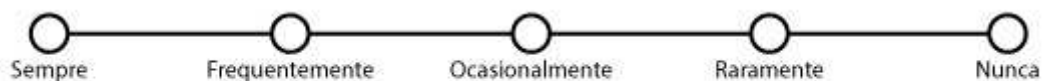
b. Assiste vídeo em outra língua, além do Português.



c. Tem contato com familiares ou pessoas próximas que se dirigem a ele/ela em uma segunda língua.



d. Tem interesse em aprender outras línguas.



Apêndice B - Folha A4 em que as crianças escolhiam o nome para o robô do experimento

[illegible]

ANEXOS

Anexo A - Palavras novas usadas na fase de teste para todos os participantes

[e]	[u]	[a]	[o]
Zanê	Pitú	Difá	Lerô
Joxê	Kezú	Xudá	Zabô
Mitê	Somú	Sebá	Vilô
rupê	nalú	gopá	gunô

Consoantes	Nomes –[e]		Nomes –[u]	
	Inicial	Medial	Inicial	Medial
[p]		1	1	
[b]				
[t]		1		1
[d]				
[k]			1	
[g]				
[m]	1			1
[n]		1		
[ɲ]				
[f]				
[v]				
[s]			1	
[z]	1			1
[ʃ]		1		
[ʒ]	1			
[x]	1			
[r]				
[l]				
[ʎ]				

Consoantes	Nomes –[a]		Nomes –[o]	
	Inicial	Medial	Inicial	Medial
[p]		1		
[b]		1		1
[t]				
[d]	1	1		
[k]				
[g]	1		1	
[m]				
[n]				1
[ɲ]				
[f]		1		
[v]			1	
[s]	1			
[z]			1	
[ʃ]	1			
[ʒ]				
[x]				
[r]				1
[l]			1	1
[ʎ]				

Vogal	Nomes –[e]	Nomes –[u]	Nomes –[a]	Nomes –[o]
[a]	1	1		1
[e]		1	1	1
[i]	1	1	1	1
[o]	1	1	1	
[u]	1		1	1

Fonte: Henrique, 2023

Anexo B - Estímulos sonoros da fase de familiarização para os participantes do grupo experimental

DETERMINANTES KOI e ZIS

[ʒi'xe]	[ʒe'tu]
[ta'e]	[mi'bu]
[bi've]	[na'du]
[do'fe]	[pɔ'fu]
[zi'ke]	[li'mu]
[nɛ'be]	[ko'zu]
[sɔ'de]	[va'u]
[fɔ'pe]	[fu'pu]
[la'ne]	[ti'gu]
[pe'ke]	[se'u]
[ka'ɛ]	[zi'nu]
[u'ɛ]	[dɔ'ɪu]
[gu'le]	[xo'vu]
[vo'me]	[u'u]
[xe'se]	[gɛ'su]
[mu'ze]	[bi'ɲu]

Consoantes	Nomes –[e]		Nomes –[u]	
	Inicial	Medial	Inicial	Medial
[p]	1	1	1	1
[b]	1	1	1	1
[t]	1		1	1
[d]	1	1	1	1
[k]	1	1	1	
[g]	1		1	1
[m]	1	1	1	1
[n]	1	1	1	1
[ɲ]		1		1
[f]	1	1	1	1
[v]	1	1	1	1
[s]	1	1	1	1
[z]	1	1	1	1
[ʃ]	1		1	1
[ʒ]	1	1	1	1
[x]	1	1	1	
[ɣ]		1		1
[l]	1	1	1	
[ʎ]		1		1

Vogal	Nomes –[e]	Nomes –[u]
[a]	3	2
[ɪ]	2	3
[e]	2	2
[i]	2	3
[ɔ]	2	2
[o]	2	2
[u]	3	2

DETERMINANTES DAI e NUR

[pi'fo]	[dɛ'ma]
[sa'vo]	[nɔ'pa]
[kɛ'lo]	[gu'a]
[de'zo]	[ku' la]
[vɔ'o]	[vi'fa]
[zu' ɲo]	[tɛ'ba]
[fo'po]	[la'za]
[a'do]	[mo ɜa]
[mɔ'go]	[pɔ'ga]
[xo'no]	[sa'ta]
[ʒɛ'to]	[bɛ'ia]
[nu'bo]	[xo' ɲa]
[le'ko]	[fi'da]
[bi'so]	[ɛ'e'na]
[ga'xo]	[zu'va]
[tɛ'lo]	[i'la]

Consoantes	Nomes –[o]		Nomes –[a]	
	Inicial	Medial	Inicial	Medial
[p]	1	1	1	1
[b]	1	1	1	1
[t]	1	1	1	1
[d]	1	1	1	1
[k]	1	1	1	
[g]	1	1	1	1
[m]	1		1	1
[n]	1	1	1	1
[ɲ]		1		1
[f]	1	1	1	1
[v]	1	1	1	1
[s]	1	1	1	
[z]	1	1	1	1
[ʃ]	1		1	1
[ç]	1		1	1
[x]	1	1	1	
[j]		1		1
[l]	1	1	1	1
[ʎ]		1		1

Vogal	Nomes –[o]	Nomes –[a]
[a]	3	2
[ɛ]	3	2
[e]	2	2
[i]	2	3
[ɔ]	2	2
[o]	2	2
[u]	2	3

Fonte: Henrique, 2023